

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2003, 10:14:14 : Search time 9019.27 Seconds
(without alignments)
11121.789 Million cell updates/sec

Title: US-09-945-326-1
Perfect score: 2452
Sequence: 1 cgtgtgtgtctcctgcggc.....actgttaaaaaaaaaa 2452

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl:*
- 1: gb.ba.*
 - 2: gb.htg.*
 - 3: gb.in.*
 - 4: gb.om.*
 - 5: gb.ov.*
 - 6: gb.pat.*
 - 7: gb.ph.*
 - 8: gb.pl.*
 - 9: gb.pr.*
 - 10: gb.ro.*
 - 11: gb.sts.*
 - 12: gb.sy.*
 - 13: gb.un.*
 - 14: gb.vi.*
 - 15: em.ba.*
 - 16: em.fun.*
 - 17: em.hum.*
 - 18: em.in.*
 - 19: em.mu.*
 - 20: em.om.*
 - 21: em.or.*
 - 22: em.ov.*
 - 23: em.pat.*
 - 24: em.ph.*
 - 25: em.pl.*
 - 26: em.ro.*
 - 27: em.sts.*
 - 28: em.un.*
 - 29: em.vi.*
 - 30: em.htg_hum.*
 - 31: em.htg_inv.*
 - 32: em.htg_other.*
 - 33: em.htg_mus.*
 - 34: em.htg_pln.*
 - 35: em.htg_rod.*
 - 36: em.htg_mam.*
 - 37: em.htg_vrt.*
 - 38: em.sy.*
 - 39: em.htgo_hum.*
 - 40: em.htgo_mus.*
 - 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2452	100.0	2452	6	AX394747 Sequence
2	2445.8	99.7	2494	9	BC007970 Homo sapi
3	2422.4	98.8	2448	9	BC013354 Homo sapi
4	2391.8	97.5	2411	9	AF327351 Homo sapi
5	2382.2	97.2	2399	6	BD158773 Primer fo
6	2382.2	97.2	2399	9	AK024012 Homo sapi
7	2115.4	86.3	2400	6	BD157796 Primer fo
8	2115.4	86.3	2400	9	AK022568 Homo sapi
9	1863	76.0	1863	6	AX394749 Sequence
10	1439	58.7	1494	9	BC001817 Sequence
11	1388.4	56.6	2129	10	BC033277 Mus muscu
12	1388.4	56.6	2148	10	BC032213 Mus muscu
13	1388.4	56.6	3071	10	BC031137 Mus muscu
14	759.6	31.0	1995	9	AK027186 Primer fo
15	750.6	30.6	771	6	BD150623 Homo sapi
16	608	24.8	17777	2	AC022993 Primer fo
17	608	24.8	185470	9	AC112484 Homo sapi
18	530.4	21.6	544	6	BD109621 EST and e
19	518.4	21.1	543	6	AX588711 Sequence
20	499.6	20.4	772	6	BD129328 Human gen
21	487.8	19.9	563	6	BD155541 Primer fo
22	487	19.9	507	6	AX588841 Sequence
23	439.2	17.9	796	6	BD147321 Primer fo
24	418.4	17.1	431	6	AX072562 Sequence
25	396.4	16.2	2058	9	AB083302 Macaca fa
26	396.4	16.2	2103	10	BTU30817 Bos taurus
27	396	16.2	2140	4	MTMVLCD
28	394.8	16.1	2171	10	BC026559 Mus muscu
29	391.6	16.0	1965	6	BD141736 IRAP bind
30	391.6	16.0	2168	9	BC020218 Homo sapi
31	391.6	16.0	2169	9	HSHVLCD
32	391.6	16.0	2219	9	HSHVLCD
33	391.6	16.0	2231	9	BC012912 Homo sapi
34	391.6	16.0	2258	9	AK056083 Homo sapi
35	391.6	16.0	2259	9	BC000399 Homo sapi
36	391.2	16.0	1840	10	AF017176 Mus muscu
37	390.4	15.9	1439	10	AY292968 Rattus no
38	388.6	15.8	2117	6	AX401908 Sequence
39	388.6	15.8	2117	10	RATVLCAD
40	380.2	15.5	434	6	AX071059 Sequence
41	372.8	15.2	171981	3	AC008344 Drosophill
42	372.8	15.2	176813	3	AC008096 Drosophill
43	372.8	15.2	176980	2	AC020290 Drosophill
44	372.8	15.2	277421	3	AE003796 Drosophill
45	372.6	15.2	412	6	AX071159 Sequence

ALIGNMENTS

RESULT 1	AX394747	Sequence 1 from Patent WO0218582.	2452 bp	DNA	linear	PAT 18-MAY-2002
LOCUS	AX394747	Sequence 1 from Patent WO0218582.				
DEFINITION	AX394747					
ACCESSION	AX394747					
VERSION	AX394747.1	GI:21065826				
KEYWORDS						
SOURCE						
ORGANISM						
		Homo sapiens (human)				
		Homo sapiens				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
		Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
REFERENCE						
AUTHORS		Meyers, R. and Hunter, J.J.				
TITLE		62112, a novel human dehydrogenase and uses thereof				
JOURNAL		Patent: WO 0218582-A 1 07-MAR-2002;				

FEATURES		Millennium Pharmaceuticals, Inc. (US)
Source		Location/Qualifiers,
		1..2452
		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
		67..1932
CDS		/note="unnamed protein product"
		/codon_start=1
		/protein_id="CAD32049.1"
		/db_xref="GI:21065827"
		/translation="MSGGGLFLRTTAAARACRLVSTANRLRLTSPVRAFAKELF
		LGKIKKVEFPPEVSDNEINOFPLGPEKFFTEEDVDSRKIDQEGKIPDTEFLK
		SLGLFGLQPEYGLGFSNTWYSLRGETIISMDGSIITVLAHQAIGLKGILACTEE
		QKALYLPKASGEHIAFCLETPASGSDAASRSRATLSEDKHYILANGSKVITNGG
		LANIFTFAKTEVDSDGSKDKITAFIVERDFGGVINGKPEDEKLGIRGNTCEVHFE
		NKIPVNTILGVDGFKVAMNLSNGSVSGVAGLLRLRIEMETAEYACRKFQNK
		RLSEFLGIKEFALMAQKAYVMESMTYLTAGMDQFPDCEIAEMKVFSSAAQW
		CYSEALQILGLGYTRDPYRILRDLRIILLIFEGTNEILRMVIALTLQHAIRLIT
		RIHELQAKSVMDTVGRRLRDLRSLGRVLDLGTNGHVHVPVLSADSKAFENTYCF
		GRIVETLLLRFRKTIIEEQVLKRVANILINLXGMTAVLSRASRSIRIGLRNHDVEL
		LANTFCVEAYLQNLFLSLQSDAYAPENLDEQIKKVSQQLLEKRAYICAPLDRTC"
BASE COUNT		616 a 616 c 685 g 535 t
ORIGIN		
Query Match		100.0%; Score 2452; DB 6; Length 2452;
Best Local Similarity		100.0%; Pred No. 0;
Matches 2452;		Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	CGTGTGTGTCTCCCTCGCGCGCTAAGAAGGGAGACTGAGGCTGAGGCTGGGGAACATCG 60
Db	1	CGTGTGTGTCTCCCTCGCGCGCTAAGAAGGGAGACTGAGGCTGAGGCTGGGGAACATCG 60
Qy	61	GGCAGCATGAGCGGCTCGCGGCTCTCCCTCGCACACAGCGGCTCGCGCTCGTGCCTGCCGG 120
Db	61	GGCAGCATGAGCGGCTCGCGGCTCTCTCCCTCGCACACAGCGGCTCGCGCTCGTGCCTGCCGG 120
Qy	121	GGTCTGTGTGTCTACCGGGAACCGCGGCTACTCGGCACACAGCGGCTCGCGCTCGTACGAGCT 180
Db	121	GGTCTGTGTGTCTACCGGGAACCGCGGCTACTCGGCACACAGCGGCTCGTACGAGCT 180
Qy	181	TTCCCCAAAGAGCTTTTCTTAGCAAAATCAAGAAGAAAGAAAGTTTCCCATTTCCAGAA 240
Db	181	TTCCCCAAAGAGCTTTTCTTAGCAAAATCAAGAAGAAAGAAAGTTTCCCATTTCCAGAA 240
Qy	241	GTTAGCAAGATGAACATTAATGAATCAATCAGTTCTTGGGACCGCGTGGAAAAATCTTC 300
Db	241	GTTAGCCNAGATGAACATTAATGAATCAATCAGTTCTTGGGACCGCGTGGAAAAATCTTC 300
Qy	301	ACTGAAGAGTGGACTCCGAAAAATTGACAGGAAGGAAATCCAGATGAACATTTG 360
Db	301	ACTGAAGAGTGGACTCCGAAAAATTGACAGGAAGGAAATCCAGATGAACATTTG 360
Qy	361	GAGAAATGAAGAGCTAGGCGCTTTTGGGCTGCAAGTCCAGAAAGAAATATGTTGGCGCTG 420
Db	361	GAGAAATGAAGAGCTAGGCGCTTTTGGGCTGCAAGTCCAGAAAGAAATATGTTGGCGCTG 420
Qy	421	GGCTTCTCCACACCATGATCTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCATC 480
Db	421	GGCTTCTCCACACCATGATCTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCATC 480
Qy	481	ACTGTGACCTGGGAGCGCACCGAGCTATTGGCCTCAAGGGGATCATCTTGGCTGGCACT 540
Db	481	ACTGTGACCTGGGAGCGCACCGAGCTATTGGCCTCAAGGGGATCATCTTGGCTGGCACT 540
Qy	541	GAGGAGCAGAAAGCCAAATCTTGCTTAACTGGCGTCCGGGGAGCAATTTGACGCTTC 600
Db	541	GAGGAGCAGAAAGCCAAATCTTGCTTAACTGGCGTCCGGGGAGCAATTTGACGCTTC 600
Qy	601	TGCCTCAGGAGCCAGCGAGTGGGAGCGATGCAGGCTCAATCCGGGAGGAGCCACACTA 660
Db	601	TGCCTCAGGAGCCAGCGAGTGGGAGCGATGCAGGCTCAATCCGGGAGGAGCCACACTA 660

Qy	661	AGTGAAGACAGAAAGCACTACATCTCAATGGCTCCAAGGCTCTGGATTACTAATGGAGGA 720
Db	661	AGTGAAGACAGAAAGCACTACATCTCAATGGCTCCAAGGCTCTGGATTACTAATGGAGGA 720
Qy	721	CTGGCCCAATATTTTACTGTGTTTGAAGACTGAGGTCGTTGATCTGTGATGATCAGT 780
Db	721	CTGGCCCAATATTTTACTGTGTTTGAAGACTGAGGTCGTTGATCTGTGATGATCAGT 780
Qy	781	AAAGACAAAATCACAGCATTCATAGTAGAAGAGACACTTTGGTGGAGTCATCTAATGGGAAA 840
Db	781	AAAGACAAAATCACAGCATTCATAGTAGAAGAGACACTTTGGTGGAGTCATCTAATGGGAAA 840
Qy	841	CCCAGAAATAATTTAGGCATTCGGGGCTCCAAACACTTTGTGAAGTCCATTTTGAACAACACC 900
Db	841	CCCAGAAATAATTTAGGCATTCGGGGCTCCAAACACTTTGTGAAGTCCATTTTGAACAACACC 900
Qy	901	AAGATACCTGTGGAATAACATCTTTGGAGAGTCGAGATGGGTTTAAAGTGGGCCATGAAC 960
Db	901	AAGATACCTGTGGAATAACATCTTTGGAGAGTCGAGATGGGTTTAAAGTGGGCCATGAAC 960
Qy	961	ATCCTCAACAGCGCGCGGTTTCAAGATGGGAGCGCTCGTGGCTGGGCTGCTCAAGAGATTG 1020
Db	961	ATCCTCAACAGCGCGCGGTTTCAAGATGGGAGCGCTCGTGGCTGGGCTGCTCAAGAGATTG 1020
Qy	1021	ATTGAATGACTGCTGAGTACGCTGCACAAAGAAACAGTTTAAACAGAGGCTCAGTGAA 1080
Db	1021	ATTGAATGACTGCTGAGTACGCTGCACAAAGAAACAGTTTAAACAGAGGCTCAGTGAA 1080
Qy	1081	TTTGGATTGATTACAGSAGAAATTTGCACTGTGCTCAGAGGCTTAGTCTATGAGAGT 1140
Db	1081	TTTGGATTGATTACAGSAGAAATTTGCACTGTGCTCAGAGGCTTAGTCTATGAGAGT 1140
Qy	1141	ATGACCTACCTCACAGCAGGAGTGTGGGCTACACAGGAGTATCCGTACGAGCGGATCTCGTGCAC 1200
Db	1141	ATGACCTACCTCACAGCAGGAGTGTGGGCTACACAGGAGTATCCGTACGAGCGGATCTCGTGCAC 1200
Qy	1201	GCAGCCATGGTGAAGGTGTTCAGCTCCGAGGCGGCTGCGAGTGTGTGAGTGAGCGGCTG 1260
Db	1201	GCAGCCATGGTGAAGGTGTTCAGCTCCGAGGCGGCTGCGAGTGTGTGAGTGAGCGGCTG 1260
Qy	1261	CAGATCCTCGGGGCTTGGGCTACACAGGAGTATCCGTACGAGCGGATCTCGTGCAC 1320
Db	1261	CAGATCCTCGGGGCTTGGGCTACACAGGAGTATCCGTACGAGCGGATCTCGTGCAC 1320
Qy	1321	ACCGCATCTCTCTATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCCCTG 1380
Db	1321	ACCGCATCTCTCTATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCCCTG 1380
Qy	1381	ACGGGTCTGAGCATGCCGGCGCATCTCTGACTACAGGATCCATGAGCTTAAACAGGCC 1440
Db	1381	ACGGGTCTGAGCATGCCGGCGCATCTCTGACTACAGGATCCATGAGCTTAAACAGGCC 1440
Qy	1441	AAAGTGAAGCAGTATGATACCGTTGGCGGAGGCTTCGGGACTTCCCTGGGCCCAACT 1500
Db	1441	AAAGTGAAGCAGTATGATACCGTTGGCGGAGGCTTCGGGACTTCCCTGGGCCCAACT 1500
Qy	1501	GTGAGCTGGGCTGACAGCAACCATGAGTTGTGACCCCATCTTGGGACAGTGC 1560
Db	1501	GTGAGCTGGGCTGACAGCAACCATGAGTTGTGACCCCATCTTGGGACAGTGC 1560
Qy	1561	AACAAGTTTGAAGAGAACACCTTACTGCTTCGGCGGAGCGGCTGGAGACACTGTGCTCCGC 1620
Db	1561	AACAAGTTTGAAGAGAACACCTTACTGCTTCGGCGGAGCGGCTGGAGACACTGTGCTCCGC 1620
Qy	1621	TTTGGCAAGACCATCATGAGGAGGAGTGTGTACTGAAGCGGTGGCCACATCTCTCATC 1680
Db	1621	TTTGGCAAGACCATCATGAGGAGGAGTGTGTACTGAAGCGGTGGCCACATCTCTCATC 1680
Qy	1681	AACCTGTATGAGTACGCGGCTGTGCTCCGGGCGGAGCGCTCCATCCGCAATTTGGGCTC 1740
Db	1681	AACCTGTATGAGTACGCGGCTGTGCTCCGGGCGGAGCGCTCCATCCGCAATTTGGGCTC 1740
Qy	1741	CGCAACCCAGCAGGAGGTTCTTCTTGGCCAAACACCTTCTTGTGCGTGAAGGCTTACTTGCAG 1800

```

|||||
1741 CGCAACACGACGAGGTTCTCTTGCCCAACACCTCTCTCGTGGAGCTTACTTCGAG 1800
QY 1801 AATCTCTTCAGGCTCTCTCAGCTGGACAAATGATCTCCAGAAAACCTTAGATGACAGATT 1860
Db 1801 AATCTCTTCAGGCTCTCTCAGCTGGACAAATGATCTCCAGAAAACCTTAGATGACAGATT 1860
QY 1861 AAGAAAGTGTCCCAAGCAGATCTCTCAGAAAGGAGCCTATATCTGTGCCCCACCTCTGGAC 1920
Db 1861 AAGAAAGTGTCCCAAGCAGATCTCTCAGAAAGGAGCCTATATCTGTGCCCCACCTCTGGAC 1920
QY 1921 AGGACATGCTGAGCAGGAGGACAGTGTCCCTGCTACCGCCGCCCTACCCATGCCCCG 1980
Db 1921 AGGACATGCTGAGCAGGAGGACAGTGTCCCTGCTACCGCCGCCCTACCCATGCCCCG 1980
QY 1981 TTGCTGGATGACTCTTACTCTTTTTCAGAAAGCTGTGGATATATCAGAGTTTAAGCCTT 2040
Db 1981 TTGCTGGATGACTCTTACTCTTTTTCAGAAAGCTGTGGATATATCAGAGTTTAAGCCTT 2040
QY 2041 TTGTTCCCGCTGTCACCTGAAGGTTGTGCGCTGGCCTGGAGAGCCTCTTCCAGGTTT 2100
Db 2041 TTGTTCCCGCTGTCACCTGAAGGTTGTGCGCTGGCCTGGAGAGCCTCTTCCAGGTTT 2100
QY 2101 TGACCTGCAGCAGTCTCTTAACAGGACCATCAGCTTCTGAACCTGAGCGGAGAGA 2160
Db 2101 TGACCTGCAGCAGTCTCTTAACAGGACCATCAGCTTCTGAACCTGAGCGGAGAGA 2160
QY 2161 GAGAATGGAATGCTGACCCCTGGAAGTGGCGGATTCCTGTGCTATGAGGAGACCAT 2220
Db 2161 GAGAATGGAATGCTGACCCCTGGAAGTGGCGGATTCCTGTGCTATGAGGAGACCAT 2220
QY 2221 AGTGGAACTGGGCTTATGCTGCTCCAGGCTGAGGTGGTGGGACCTGTGTC 2280
Db 2221 AGTGGAACTGGGCTTATGCTGCTCCAGGCTGAGGTGGTGGGACCTGTGTC 2280
QY 2281 AGTGTGGATAGCCATTTCTGCTCAACACACATCTCTTAAGAACACCTTGAAGCTCT 2340
Db 2281 AGTGTGGATAGCCATTTCTGCTCAACACACATCTCTTAAGAACACCTTGAAGCTCT 2340
QY 2341 GTCTGGTCATTCATTAACTAGAGCAGAGGACCTTAAACATGTACCAGGACCAT 2400
Db 2341 GTCTGGTCATTCATTAACTAGAGCAGAGGACCTTAAACATGTACCAGGACCAT 2400
QY 2401 TAACAAAGATATAAATGTCACATCTGTGTACTGTTAAAAAANAANA 2452
Db 2401 TAACAAAGATATAAATGTCACATCTGTGTACTGTTAAAAAANAANA 2452

```

FEATURES

source

1. .2494

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14452 IMAGE:4304209"

/tissue_type="uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_46"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .2494

/gene="VLCAD"

/note="synonyms: NP0002, MGC14452, ACAD-9"

/db_xref="LocusID:28976"

89. .1954

/codon_start=1

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

/protein_id="AAH07970.1"

/db_xref="GI:14044102"

/db_xref="LocusID:28976"

/translation="MSGCGFLRTTAAACRGLVSTANRRLLRTSPVPVAFKELF

LGKIKKEVPFPEVSQDEINEINQFLGPEKFTTEVDSRKTIDQSGKIPDETLEKLK

SLGLFLQVPEEYGGAGFSTNTSRIGETISMDGTSITVTLAAHOATGLKGIILAGTEE

QKARYLPKLAGEHIAFCLTEPASGDAASIRSRATLSTLSDKHYILNGSKVWTNGG

LANITVFATKEVVDGSKVKITAFIVERDFFGGVTKGPKDKLGIRSNCEVHFE

NKTPVBNILGEVDGFKVAMNLTNSRFSMSVAGVLKRLTIETACTRKQFNK

RLSEFGLIOEKFAKMAQKAVVMESMTYLTAGMLDQPGFDCSIEAMVVFSSAAWQ

CVSEALQILGLGTYTDPYERLRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

RHELQAKVSTVNDTVGRRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

gene

CDS

Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 2494)

Strausberg, R.

Direct Submission

Submitted (11-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-re@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLML)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., McDowell, J., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., Maduro, Q.L., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLML at: <http://lmage.llnl.gov>

Series: IRAL Plate: 20 Row: P Column: 20

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

source

1. .2494

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14452 IMAGE:4304209"

/tissue_type="uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_46"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .2494

/gene="VLCAD"

/note="synonyms: NP0002, MGC14452, ACAD-9"

/db_xref="LocusID:28976"

89. .1954

/codon_start=1

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

/protein_id="AAH07970.1"

/db_xref="GI:14044102"

/db_xref="LocusID:28976"

/translation="MSGCGFLRTTAAACRGLVSTANRRLLRTSPVPVAFKELF

LGKIKKEVPFPEVSQDEINEINQFLGPEKFTTEVDSRKTIDQSGKIPDETLEKLK

SLGLFLQVPEEYGGAGFSTNTSRIGETISMDGTSITVTLAAHOATGLKGIILAGTEE

QKARYLPKLAGEHIAFCLTEPASGDAASIRSRATLSTLSDKHYILNGSKVWTNGG

LANITVFATKEVVDGSKVKITAFIVERDFFGGVTKGPKDKLGIRSNCEVHFE

NKTPVBNILGEVDGFKVAMNLTNSRFSMSVAGVLKRLTIETACTRKQFNK

RLSEFGLIOEKFAKMAQKAVVMESMTYLTAGMLDQPGFDCSIEAMVVFSSAAWQ

CVSEALQILGLGTYTDPYERLRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

RHELQAKVSTVNDTVGRRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

1. .2494

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14452 IMAGE:4304209"

/tissue_type="uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_46"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .2494

/gene="VLCAD"

/note="synonyms: NP0002, MGC14452, ACAD-9"

/db_xref="LocusID:28976"

89. .1954

/codon_start=1

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

/protein_id="AAH07970.1"

/db_xref="GI:14044102"

/db_xref="LocusID:28976"

/translation="MSGCGFLRTTAAACRGLVSTANRRLLRTSPVPVAFKELF

LGKIKKEVPFPEVSQDEINEINQFLGPEKFTTEVDSRKTIDQSGKIPDETLEKLK

SLGLFLQVPEEYGGAGFSTNTSRIGETISMDGTSITVTLAAHOATGLKGIILAGTEE

QKARYLPKLAGEHIAFCLTEPASGDAASIRSRATLSTLSDKHYILNGSKVWTNGG

LANITVFATKEVVDGSKVKITAFIVERDFFGGVTKGPKDKLGIRSNCEVHFE

NKTPVBNILGEVDGFKVAMNLTNSRFSMSVAGVLKRLTIETACTRKQFNK

RLSEFGLIOEKFAKMAQKAVVMESMTYLTAGMLDQPGFDCSIEAMVVFSSAAWQ

CVSEALQILGLGTYTDPYERLRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

RHELQAKVSTVNDTVGRRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

1. .2494

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14452 IMAGE:4304209"

/tissue_type="uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_46"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .2494

/gene="VLCAD"

/note="synonyms: NP0002, MGC14452, ACAD-9"

/db_xref="LocusID:28976"

89. .1954

/codon_start=1

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

/protein_id="AAH07970.1"

/db_xref="GI:14044102"

/db_xref="LocusID:28976"

/translation="MSGCGFLRTTAAACRGLVSTANRRLLRTSPVPVAFKELF

LGKIKKEVPFPEVSQDEINEINQFLGPEKFTTEVDSRKTIDQSGKIPDETLEKLK

SLGLFLQVPEEYGGAGFSTNTSRIGETISMDGTSITVTLAAHOATGLKGIILAGTEE

QKARYLPKLAGEHIAFCLTEPASGDAASIRSRATLSTLSDKHYILNGSKVWTNGG

LANITVFATKEVVDGSKVKITAFIVERDFFGGVTKGPKDKLGIRSNCEVHFE

NKTPVBNILGEVDGFKVAMNLTNSRFSMSVAGVLKRLTIETACTRKQFNK

RLSEFGLIOEKFAKMAQKAVVMESMTYLTAGMLDQPGFDCSIEAMVVFSSAAWQ

CVSEALQILGLGTYTDPYERLRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

RHELQAKVSTVNDTVGRRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

1. .2494

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14452 IMAGE:4304209"

/tissue_type="uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_46"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .2494

/gene="VLCAD"

/note="synonyms: NP0002, MGC14452, ACAD-9"

/db_xref="LocusID:28976"

89. .1954

/codon_start=1

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

/protein_id="AAH07970.1"

/db_xref="GI:14044102"

/db_xref="LocusID:28976"

/translation="MSGCGFLRTTAAACRGLVSTANRRLLRTSPVPVAFKELF

LGKIKKEVPFPEVSQDEINEINQFLGPEKFTTEVDSRKTIDQSGKIPDETLEKLK

SLGLFLQVPEEYGGAGFSTNTSRIGETISMDGTSITVTLAAHOATGLKGIILAGTEE

QKARYLPKLAGEHIAFCLTEPASGDAASIRSRATLSTLSDKHYILNGSKVWTNGG

LANITVFATKEVVDGSKVKITAFIVERDFFGGVTKGPKDKLGIRSNCEVHFE

NKTPVBNILGEVDGFKVAMNLTNSRFSMSVAGVLKRLTIETACTRKQFNK

RLSEFGLIOEKFAKMAQKAVVMESMTYLTAGMLDQPGFDCSIEAMVVFSSAAWQ

CVSEALQILGLGTYTDPYERLRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

RHELQAKVSTVNDTVGRRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

1. .2494

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14452 IMAGE:4304209"

/tissue_type="uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_46"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .2494

/gene="VLCAD"

/note="synonyms: NP0002, MGC14452, ACAD-9"

/db_xref="LocusID:28976"

89. .1954

/codon_start=1

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

/protein_id="AAH07970.1"

/db_xref="GI:14044102"

/db_xref="LocusID:28976"

/translation="MSGCGFLRTTAAACRGLVSTANRRLLRTSPVPVAFKELF

LGKIKKEVPFPEVSQDEINEINQFLGPEKFTTEVDSRKTIDQSGKIPDETLEKLK

SLGLFLQVPEEYGGAGFSTNTSRIGETISMDGTSITVTLAAHOATGLKGIILAGTEE

QKARYLPKLAGEHIAFCLTEPASGDAASIRSRATLSTLSDKHYILNGSKVWTNGG

LANITVFATKEVVDGSKVKITAFIVERDFFGGVTKGPKDKLGIRSNCEVHFE

NKTPVBNILGEVDGFKVAMNLTNSRFSMSVAGVLKRLTIETACTRKQFNK

RLSEFGLIOEKFAKMAQKAVVMESMTYLTAGMLDQPGFDCSIEAMVVFSSAAWQ

CVSEALQILGLGTYTDPYERLRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

RHELQAKVSTVNDTVGRRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

1. .2494

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14452 IMAGE:4304209"

/tissue_type="uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_46"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .2494

/gene="VLCAD"

/note="synonyms: NP0002, MGC14452, ACAD-9"

/db_xref="LocusID:28976"

89. .1954

/codon_start=1

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

/protein_id="AAH07970.1"

/db_xref="GI:14044102"

/db_xref="LocusID:28976"

/translation="MSGCGFLRTTAAACRGLVSTANRRLLRTSPVPVAFKELF

LGKIKKEVPFPEVSQDEINEINQFLGPEKFTTEVDSRKTIDQSGKIPDETLEKLK

SLGLFLQVPEEYGGAGFSTNTSRIGETISMDGTSITVTLAAHOATGLKGIILAGTEE

QKARYLPKLAGEHIAFCLTEPASGDAASIRSRATLSTLSDKHYILNGSKVWTNGG

LANITVFATKEVVDGSKVKITAFIVERDFFGGVTKGPKDKLGIRSNCEVHFE

NKTPVBNILGEVDGFKVAMNLTNSRFSMSVAGVLKRLTIETACTRKQFNK

RLSEFGLIOEKFAKMAQKAVVMESMTYLTAGMLDQPGFDCSIEAMVVFSSAAWQ

CVSEALQILGLGTYTDPYERLRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

RHELQAKVSTVNDTVGRRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

1. .2494

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14452 IMAGE:4304209"

/tissue_type="uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_46"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .2494

/gene="VLCAD"

/note="synonyms: NP0002, MGC14452, ACAD-9"

/db_xref="LocusID:28976"

89. .1954

/codon_start=1

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

/protein_id="AAH07970.1"

/db_xref="GI:14044102"

/db_xref="LocusID:28976"

/translation="MSGCGFLRTTAAACRGLVSTANRRLLRTSPVPVAFKELF

LGKIKKEVPFPEVSQDEINEINQFLGPEKFTTEVDSRKTIDQSGKIPDETLEKLK

SLGLFLQVPEEYGGAGFSTNTSRIGETISMDGTSITVTLAAHOATGLKGIILAGTEE

QKARYLPKLAGEHIAFCLTEPASGDAASIRSRATLSTLSDKHYILNGSKVWTNGG

LANITVFATKEVVDGSKVKITAFIVERDFFGGVTKGPKDKLGIRSNCEVHFE

NKTPVBNILGEVDGFKVAMNLTNSRFSMSVAGVLKRLTIETACTRKQFNK

RLSEFGLIOEKFAKMAQKAVVMESMTYLTAGMLDQPGFDCSIEAMVVFSSAAWQ

CVSEALQILGLGTYTDPYERLRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

RHELQAKVSTVNDTVGRRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

1. .2494

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14452 IMAGE:4304209"

/tissue_type="uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_46"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .2494

/gene="VLCAD"

/note="synonyms: NP0002, MGC14452, ACAD-9"

/db_xref="LocusID:28976"

89. .1954

/codon_start=1

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

/protein_id="AAH07970.1"

/db_xref="GI:14044102"

/db_xref="LocusID:28976"

/translation="MSGCGFLRTTAAACRGLVSTANRRLLRTSPVPVAFKELF

LGKIKKEVPFPEVSQDEINEINQFLGPEKFTTEVDSRKTIDQSGKIPDETLEKLK

SLGLFLQVPEEYGGAGFSTNTSRIGETISMDGTSITVTLAAHOATGLKGIILAGTEE

QKARYLPKLAGEHIAFCLTEPASGDAASIRSRATLSTLSDKHYILNGSKVWTNGG

LANITVFATKEVVDGSKVKITAFIVERDFFGGVTKGPKDKLGIRSNCEVHFE

NKTPVBNILGEVDGFKVAMNLTNSRFSMSVAGVLKRLTIETACTRKQFNK

RLSEFGLIOEKFAKMAQKAVVMESMTYLTAGMLDQPGFDCSIEAMVVFSSAAWQ

CVSEALQILGLGTYTDPYERLRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

RHELQAKVSTVNDTVGRRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

1. .2494

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14452 IMAGE:4304209"

/tissue_type="uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_46"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .2494

/gene="VLCAD"

/note="synonyms: NP0002, MGC14452, ACAD-9"

/db_xref="LocusID:28976"

89. .1954

/codon_start=1

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

/protein_id="AAH07970.1"

/db_xref="GI:14044102"

/db_xref="LocusID:28976"

/translation="MSGCGFLRTTAAACRGLVSTANRRLLRTSPVPVAFKELF

LGKIKKEVPFPEVSQDEINEINQFLGPEKFTTEVDSRKTIDQSGKIPDETLEKLK

SLGLFLQVPEEYGGAGFSTNTSRIGETISMDGTSITVTLAAHOATGLKGIILAGTEE

QKARYLPKLAGEHIAFCLTEPASGDAASIRSRATLSTLSDKHYILNGSKVWTNGG

LANITVFATKEVVDGSKVKITAFIVERDFFGGVTKGPKDKLGIRSNCEVHFE

NKTPVBNILGEVDGFKVAMNLTNSRFSMSVAGVLKRLTIETACTRKQFNK

RLSEFGLIOEKFAKMAQKAVVMESMTYLTAGMLDQPGFDCSIEAMVVFSSAAWQ

CVSEALQILGLGTYTDPYERLRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

RHELQAKVSTVNDTVGRRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

1. .2494

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14452 IMAGE:4304209"

/tissue_type="uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_46"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .2494

/gene="VLCAD"

/note="synonyms: NP0002, MGC14452, ACAD-9"

/db_xref="LocusID:28976"

89. .1954

/codon_start=1

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

/protein_id="AAH07970.1"

/db_xref="GI:14044102"

/db_xref="LocusID:28976"

/translation="MSGCGFLRTTAAACRGLVSTANRRLLRTSPVPVAFKELF

LGKIKKEVPFPEVSQDEINEINQFLGPEKFTTEVDSRKTIDQSGKIPDETLEKLK

SLGLFLQVPEEYGGAGFSTNTSRIGETISMDGTSITVTLAAHOATGLKGIILAGTEE

QKARYLPKLAGEHIAFCLTEPASGDAASIRSRATLSTLSDKHYILNGSKVWTNGG

LANITVFATKEVVDGSKVKITAFIVERDFFGGVTKGPKDKLGIRSNCEVHFE

NKTPVBNILGEVDGFKVAMNLTNSRFSMSVAGVLKRLTIETACTRKQFNK

RLSEFGLIOEKFAKMAQKAVVMESMTYLTAGMLDQPGFDCSIEAMVVFSSAAWQ

CVSEALQILGLGTYTDPYERLRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

RHELQAKVSTVNDTVGRRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

1. .2494

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14452 IMAGE:4304209"

/tissue_type="uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_46"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .2494

/gene="VLCAD"

/note="synonyms: NP0002, MGC14452, ACAD-9"

/db_xref="LocusID:28976"

89. .1954

/codon_start=1

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

/protein_id="AAH07970.1"

/db_xref="GI:14044102"

/db_xref="LocusID:28976"

/translation="MSGCGFLRTTAAACRGLVSTANRRLLRTSPVPVAFKELF

LGKIKKEVPFPEVSQDEINEINQFLGPEKFTTEVDSRKTIDQSGKIPDETLEKLK

SLGLFLQVPEEYGGAGFSTNTSRIGETISMDGTSITVTLAAHOATGLKGIILAGTEE

QKARYLPKLAGEHIAFCLTEPASGDAASIRSRATLSTLSDKHYILNGSKVWTNGG

LANITVFATKEVVDGSKVKITAFIVERDFFGGVTKGPKDKLGIRSNCEVHFE

NKTPVBNILGEVDGFKVAMNLTNSRFSMSVAGVLKRLTIETACTRKQFNK

RLSEFGLIOEKFAKMAQKAVVMESMTYLTAGMLDQPGFDCSIEAMVVFSSAAWQ

CVSEALQILGLGTYTDPYERLRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

RHELQAKVSTVNDTVGRRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

1. .2494

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14452 IMAGE:4304209"

/tissue_type="uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_46"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .2494

/gene="VLCAD"

/note="synonyms: NP0002, MGC14452, ACAD-9"

/db_xref="LocusID:28976"

89. .1954

/codon_start=1

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

/protein_id="AAH07970.1"

/db_xref="GI:14044102"

/db_xref="LocusID:28976"

/translation="MSGCGFLRTTAAACRGLVSTANRRLLRTSPVPVAFKELF

LGKIKKEVPFPEVSQDEINEINQFLGPEKFTTEVDSRKTIDQSGKIPDETLEKLK

SLGLFLQVPEEYGGAGFSTNTSRIGETISMDGTSITVTLAAHOATGLKGIILAGTEE

QKARYLPKLAGEHIAFCLTEPASGDAASIRSRATLSTLSDKHYILNGSKVWTNGG

LANITVFATKEVVDGSKVKITAFIVERDFFGGVTKGPKDKLGIRSNCEVHFE

NKTPVBNILGEVDGFKVAMNLTNSRFSMSVAGVLKRLTIETACTRKQFNK

RLSEFGLIOEKFAKMAQKAVVMESMTYLTAGMLDQPGFDCSIEAMVVFSSAAWQ

CVSEALQILGLGTYTDPYERLRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

RHELQAKVSTVNDTVGRRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

1. .2494

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14452 IMAGE:4304209"

/tissue_type="uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_46"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .2494

/gene="VLCAD"

/note="synonyms: NP0002, MGC14452, ACAD-9"

/db_xref="LocusID:28976"

89. .1954

/codon_start=1

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

/protein_id="AAH07970.1"

/db_xref="GI:14044102"

/db_xref="LocusID:28976"

/translation="MSGCGFLRTTAAACRGLVSTANRRLLRTSPVPVAFKELF

LGKIKKEVPFPEVSQDEINEINQFLGPEKFTTEVDSRKTIDQSGKIPDETLEKLK

SLGLFLQVPEEYGGAGFSTNTSRIGETISMDGTSITVTLAAHOATGLKGIILAGTEE

QKARYLPKLAGEHIAFCLTEPASGDAASIRSRATLSTLSDKHYILNGSKVWTNGG

LANITVFATKEVVDGSKVKITAFIVERDFFGGVTKGPKDKLGIRSNCEVHFE

NKTPVBNILGEVDGFKVAMNLTNSRFSMSVAGVLKRLTIETACTRKQFNK

RLSEFGLIOEKFAKMAQKAVVMESMTYLTAGMLDQPGFDCSIEAMVVFSSAAWQ

CVSEALQILGLGTYTDPYERLRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

RHELQAKVSTVNDTVGRRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

1. .2494

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14452 IMAGE:4304209"

/tissue_type="uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_46"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .2494

/gene="VLCAD"

/note="synonyms: NP0002, MGC14452, ACAD-9"

/db_xref="LocusID:28976"

89. .1954

/codon_start=1

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

/protein_id="AAH07970.1"

/db_xref="GI:14044102"

/db_xref="LocusID:28976"

/translation="MSGCGFLRTTAAACRGLVSTANRRLLRTSPVPVAFKELF

LGKIKKEVPFPEVSQDEINEINQFLGPEKFTTEVDSRKTIDQSGKIPDETLEKLK

SLGLFLQVPEEYGGAGFSTNTSRIGETISMDGTSITVTLAAHOATGLKGIILAGTEE

QKARYLPKLAGEHIAFCLTEPASGDAASIRSRATLSTLSDKHYILNGSKVWTNGG

LANITVFATKEVVDGSKVKITAFIVERDFFGGVTKGPKDKLGIRSNCEVHFE

NKTPVBNILGEVDGFKVAMNLTNSRFSMSVAGVLKRLTIETACTRKQFNK

RLSEFGLIOEKFAKMAQKAVVMESMTYLTAGMLDQPGFDCSIEAMVVFSSAAWQ

CVSEALQILGLGTYTDPYERLRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

RHELQAKVSTVNDTVGRRLSLRTDGLTGNHGVHPVSLADSNKFPENTYCF

1. .2494

Tue Sep 9 10:14:35 2003

BASE COUNT		643 a		621 c		692 g		538 t			
ORIGIN											
Query Match		99.7%;		Score 2445.8;		DB 9;		Length 2494;			
Best Local Similarity		99.9%;		Pred. No. 0;							
Matches 2447;		Conservative		0; Mismatches		2; Indels		0; Gaps			
GRTVETLLLRFGKTIMEQVLKRVANILINLYGWTAVLSRASRSIRIGLRNHDHVEL											
LANTFCVAYLQNLFSLSQLDKYAPENLDEQIKKVSOQILEKRAYICAHPLDRFC											
4 GTGTGTGTCCTCGCGCGCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGGAACTCGGGC		63		621		692		538			
26 GTGTGTGTCCTCGCGCGCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGGAACTCGGGC		85									
64 AGCATGAGGGGTGCGGGCTCTTCTGCGCCACACCGGTGGGCTGTCGCTGCGGGGT		123									
86 AGCATGAGGGGTGCGGGCTCTTCTGCGCCACACCGGTGGGCTGTCGCTGCGGGGT		145									
124 CTGGTGTCTCTACCCGGAACCGGGCTGCTGCGCCACACCGGTGGGCTGTCGCTGCGGGGT		183									
146 CTGGTGTCTCTACCCGGAACCGGGCTGCTGCGCCACACCGGTGGGCTGTCGCTGCGGGGT		205									
184 GCCAAGAGGCTTTTCTAGSCAAAATCAAGAGAGAAAGTTTCCCATTTCCAGAGTT		243									
206 GCCAAGAGGCTTTTCTAGSCAAAATCAAGAGAGAAAGTTTCCCATTTCCAGAGTT		265									
244 AGCCAGATGAATTAATGAATCAATCAGTTCTTGGGACCCGCTGACGAGCTTTC		303									
266 AGCCAGATGAATTAATGAATCAATCAGTTCTTGGGACCCGCTGACGAGCTTTC		325									
304 GAAGAGGTGACTCCCGAAAATTTGACAGAGAGGAGAAATCCGATGAAGTTGGAG		363									
326 GAAGAGGTGACTCCCGAAAATTTGACAGAGAGGAGAAATCCGATGAAGTTGGAG		385									
364 AAATGAAGAGCTTAGGGCTTTTGGGCTCAAGTCCCAAGAAATATGGTGGCTGGGC		423									
386 AAATGAAGAGCTTAGGGCTTTTGGGCTCAAGTCCCAAGAAATATGGTGGCTGGGC		445									
424 TTCTCCAAACCATGTACTCAAGACTAGGGGAGATCATCAGATGAGTGGGTCCATCACT		483									
446 TTCTCCAAACCATGTACTCAAGACTAGGGGAGATCATCAGATGAGTGGGTCCATCACT		505									
484 GTGACCTGGGAGGCGCAGGCTATTGGCTCAAGGGATCATCTTGGCTGGCAGCTGAG		543									
506 GTGACCTGGGAGGCGCAGGCTATTGGCTCAAGGGATCATCTTGGCTGGCAGCTGAG		565									
544 GAGCAGAAAGCAAAATCTTCCCTAAACTGGCGTCCGCGGAGCAGATGAGCGCTTTCG		603									
566 GAGCAGAAAGCAAAATCTTCCCTAAACTGGCGTCCGCGGAGCAGATGAGCGCTTTCG		625									
604 CTCAGGAGGCGCAGGCTAGGAGGATGAGGCTCAATCCGAGCAGAGCCACACTAAGT		663									
626 CTCAGGAGGCGCAGGCTAGGAGGATGAGGCTCAATCCGAGCAGAGCCACACTAAGT		685									
664 GAAGCAGAAAGCAAAATCTTCCCTAAACTGGCGTCCGCGGAGCAGATGAGCGCTTTCG		723									
686 GAAGCAGAAAGCAAAATCTTCCCTAAACTGGCGTCCGCGGAGCAGATGAGCGCTTTCG		745									
724 GCCAATATTTTACTGTGTTTCCAAAGACTGAGGCTGTTGATTTCTGATGATCAGTGAAA		783									
746 GCCAATATTTTACTGTGTTTCCAAAGACTGAGGCTGTTGATTTCTGATGATCAGTGAAA		805									
784 GACAAATCACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCC		843									
806 GACAAATCACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCC		865									
844 GAAGATPAATTAAGGATTCGGGGCTCCCAACACTTGTGAGTCCATTTTGAAGAACCAAG		903									
866 GAAGATPAATTAAGGATTCGGGGCTCCCAACACTTGTGAGTCCATTTTGAAGAACCAAG		925									
904 ATACCTGTGAAACATCTTGGAGAGTCCGAGATGGGTTTAAAGTGGCCATGAACATC		963									
926 ATACCTGTGAAACATCTTGGAGAGTCCGAGATGGGTTTAAAGTGGCCATGAACATC		985									

964	CTCAACAGCGCGCGGTTACAGATGGCGAGCGTCTGGCTGGGCTGCTCAAGAGATTGATT	1023
986	CTCAACAGCGCGCGGTTACAGATGGCGAGCGTCTGGCTGGGCTGCTCAAGAGATTGATT	1045
1024	GAATGACTGCTGAGTACCGCTGCACAAGGAAACAGTTTAAACAAGAGGCTCAGTGAATTT	1083
1046	GAATGACTGCTGAGTACCGCTGCACAAGGAAACAGTTTAAACAAGAGGCTCAGTGAATTT	1105
1084	GGATTGATTGAGGAGAAATTTGCACCTGATGGCTCAGAAGGCTTACGTATGAGAGATG	1143
1106	GGATTGATTGAGGAGAAATTTGCACCTGATGGCTCAGAAGGCTTACGTATGAGAGATG	1165
1144	ACCTACCTCACAGAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGCA	1203
1166	ACCTACCTCACAGAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGCA	1225
1204	GCCATGTGAAGTGTTCAGCTCCGAGCGCCCTGGCAGTGTGTAGTGAGGCGCTGCAG	1263
1226	GCCATGTGAAGTGTTCAGCTCCGAGCGCCCTGGCAGTGTGTAGTGAGGCGCTGCAG	1285
1264	ATCCTCGGGGCTTGGGCTACACAGGAGTATCCGTACCGAGCGCATCTGCTGACACC	1323
1286	ATCCTCGGGGCTTGGGCTACACAGGAGTATCCGTACCGAGCGCATCTGCTGACACC	1345
1324	CGCATCTCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCCGTGACG	1383
1346	CGCATCTCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCCGTGACG	1405
1384	GGTCTGAGCATGCGGCGGCATCTGACTACCAAGGATCCATGAGCTTTAAACAGGCCAAA	1443
1406	GGTCTGAGCATGCGGCGGCATCTGACTACCAAGGATCCATGAGCTTTAAACAGGCCAAA	1465
1444	GTGACAGCATGATGATACCGTTGGCGGAGGCTTGGGACTCCCTGGGCGCAACTGTG	1503
1466	GTGACAGCATGATGATACCGTTGGCGGAGGCTTGGGACTCCCTGGGCGCAACTGTG	1525
1504	GACCTGGGCTGACAGGCAACCATGAGTGTGACCCCTCTGCGGAGCTGCGGAGCTGCAAC	1563
1526	GACCTGGGCTGACAGGCAACCATGAGTGTGACCCCTCTGCGGAGCTGCGGAGCTGCAAC	1585
1564	AGTTTGGAGGAACACCTACTGCTTCGCGGAGGCTTGGGACTCCCTGGGCGCAACTGTG	1623
1586	AGTTTGGAGGAACACCTACTGCTTCGCGGAGGCTTGGGACTCCCTGGGCGCAACTGTG	1645
1624	GGCAAGACCATCATGGAGGAGCAGCTGCTGCTGCGGAGGCTTGGGACTCCCTGGGCGCAACTGTG	1683
1646	GGCAAGACCATCATGGAGGAGCAGCTGCTGCTGCGGAGGCTTGGGACTCCCTGGGCGCAACTGTG	1705
1684	CTGTATGGCATGACGCGCTGCTGCTGCGGAGGCTTGGGACTCCCTGGGCGCAACTGTG	1743
1706	CTGTATGGCATGACGCGCTGCTGCTGCGGAGGCTTGGGACTCCCTGGGCGCAACTGTG	1765
1744	AACCAAGACCATCATGGAGGAGCAGCTGCTGCTGCGGAGGCTTGGGACTCCCTGGGCGCAACTGTG	1803
1766	AACCAAGACCATCATGGAGGAGCAGCTGCTGCTGCGGAGGCTTGGGACTCCCTGGGCGCAACTGTG	1825
1804	CTCTTACGCTCTCTCAGCTGGACAAAGTATGCTTCCAGAAACCTAGATGAGCAGATTAAG	1863
1826	CTCTTACGCTCTCTCAGCTGGACAAAGTATGCTTCCAGAAACCTAGATGAGCAGATTAAG	1885
1864	AAAGTGTCCAGCAGATTCCTTGGGAGGAGGCTTATCTGTCGCCACCTCTGGACAGG	1923
1886	AAAGTGTCCAGCAGATTCCTTGGGAGGAGGCTTATCTGTCGCCACCTCTGGACAGG	1945
1924	ACATGTGAGGAGGAGGAGGCTGCTCCCTGCTACCGCGCGCCCTACCCATGGCGCGTTG	1983
1946	ACATGTGAGGAGGAGGAGGCTGCTCCCTGCTACCGCGCGCCCTACCCATGGCGCGTTG	2005
1984	CTGGATGACTGTTACTCTTTTTCAGAGGCTGTTGGATTTATCACAGGTTAAGCCTTTG	2043
2006	CTGGATGACTGTTACTCTTTTTCAGAGGCTGTTGGATTTATCACAGGTTAAGCCTTTG	2065
2044	TTCCCCGCTCTGCACCTGAAGGTTGTCGCTGGGCTGGGAGAGCGCTCTTCCAGGTTTGA	2103

Db 2066 TTCCCGCTCTGACCTGAAGGTTGTGCGCTGGGAGAGCTCTTCCAGGTTTGA 2125
QY 2104 CTTGAGCAGCTGCTCTTACAGGACCATCAGAGTCTTGAAGTCTGAGCGGAGAGAG 2163
Db 2126 CTTGAGCAGCTGCTCTTACAGGACCATCAGAGTCTTGAAGTCTGAGCGGAGAGAG 2185
QY 2164 AATGGAATTCGTGACCTGGAAGTCTGAGCGGATTCGTCATTGAGGACACCATAGT 2223
Db 2186 AATGGAATTCGTGACCTGGAAGTCTGAGCGGATTCGTCATTGAGGACACCATAGT 2245
QY 2224 GGAAGTGGGCTTATGCTGCTCCAGGCTGAGGCTGAGGCTGAGGCTGTCAGG 2283
Db 2246 GGAAGTGGGCTTATGCTGCTCCAGGCTGAGGCTGAGGCTGAGGCTGTCAGG 2305
QY 2284 TGTGATGAGCATTTCTGCTCAACACACATTCCTTAAGAAACAGCTTGAAGCTCTGTC 2343
Db 2306 TGTGATGAGCATTTCTGCTCAACACACATTCCTTAAGAAACAGCTTGAAGCTCTGTC 2365
QY 2344 TGGGTCTTCTTAACTAGAGCAGGACCTTAAACATGATACAGGACCATTTAA 2403
Db 2366 TGGGTCTTCTTAACTAGAGCAGGACCTTAAACATGATACAGGACCATTTAA 2425
QY 2404 CAAAGATATAAATGTCACATCTGCTGCTTAAACATGATACAGGACCATTTAA 2452
Db 2426 CAAAGATATAAATGTCACATCTGCTGCTTAAACATGATACAGGACCATTTAA 2474

RESULT 3
LOCUS BC013354
DEFINITION Homo sapiens very-long-chain acyl-CoA dehydrogenase-VLCAD, mRNA
ACCESSION BC013354
VERSION 1
KEYWORDS MGC:14970 IMAGE:3935925, complete cds.
SOURCE MGC.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2448)
Strausberg, R.D., Collins, P., Wagner, L., Shenmen, C.M., Schuler, G.D., Altshul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whitling, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Smalinski, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Warra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 2448)
Strausberg, R.
Direct Submission
Submitted (31-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk

Email: cgapbs-re@mail.nih.gov
Tissue procurement: DCTD/Dnp
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
Contact: amadan@systemsbio.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Place: 25 Row: c Column: 21.

FEATURES		Location/Qualifiers	
source		1..2448	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="MGC:14970 IMAGE:3935925"	
		/tissue_type="Lung, small cell carcinoma"	
		/clone_lib="NIH_MGC_7"	
		/lab_host="DH10B-R"	
		/note="Vector: pOTB7"	
gene	1..2448		
	/gene="VLCAD"		
	/note="synonyms: NP0002, MGC14452, ACAD-9"		
CDS	/db_xref="LocusID:28976"		
	66..1931		
	/codon_start=1		
	/product="very-long-chain acyl-CoA dehydrogenase VLCAD"		
	/protein_id="AAH13354.1"		
	/db_xref="GI:15426498"		
	/db_xref="LocusID:28976"		
	/translation="MSGCLFLRTTAAACRGLVSTANRLRLTSPVRAFAKELF LGKIKKKEVFPPEVQDELINQFLGPVEKFTFEVDSRKIDQSGKIPDETLEKLK SLGLQVLPVEEYGLFNTMYSLRGLIISMDGSIYTLAAHQAIGLGLAGTLEE OKAKYLPKLAGSHEIAAFCLTEPASGSDAASIRSRATLSEDKKHLYLNGSKVYITNGE LANITFVEAKTEVYDSDGSKITAFIVERDPGVNGTNGKEDKHYLNGSKVYITNGE NTKIPVENIIGEVGDFKVMNINLRSFMSVAGLLKRLIETAEYAKYRKQFNFK CVSEALQIGLGLYTRDYPYRILRLITRILIPETNGEILRMVIALTQLOHAGRLIT RHELKQAKYSTVMDTVGRRLRDLGTLGTHGVHVPSPADSANKFRENTRYCF GRVETLLEFRFGTITMEQSLDLKRVANILINLYGMTAVLSRASRIRIGLRNHDHEVL LANTFCEVAYLQNLFSQLDKYAPENLDEQIKKVSQQLLEKRAYICAHPLDRTCT"		
BASE COUNT	624 a 617 c 682 g 525 t		
ORIGIN			
Query Match	98.8%;	Score 2422.4;	DB 9; Length 2448;
Best Local Similarity	99.5%;	Pred. No. 0;	
Matches 2429;	Conservative 0;	Mismatches 11;	Indels 0; Gaps 0;
QY	10	GTCCCTGCGGCGCTAAGAGGGGAGACTGAGCTGAGGCTGGGGAACATCGGGCAGCATG	69
Db	9	GTCCCTGCGGCGCTAAGAGGGGAGACTGAGCTGAGGCTGGGGAACATCGGGCAGCATG	68
QY	70	ACGGCTGCGGCGCTCTTCTTCCGCGCACACCGCTGCGGCTCGCTGCGGGGCTCTGGTG	129
Db	69	ACGGCTGCGGCGCTCTTCTTCCGCGCACACCGCTGCGGCTCGCTGCGGGGCTCTGGTG	128
QY	130	GTCTCTACCGCGAACCAGCGGCTACTGCGCACACCGCGCTGTACGAGCTTTCGCCAAA	189
Db	129	GTCTCTACCGCGAACCAGCGGCTACTGCGCACACCGCGCTGTACGAGCTTTCGCCAAA	188
QY	190	GAGCTTTTCTAGGCAAAATCAAGAAGAAAGTTTCCCATTTCCAGAGTTTAGCCAA	249
Db	189	GAGCTTTTCTAGGCAAAATCAAGAAGAAAGTTTCCCATTTCCAGAGTTTAGCCAA	248
QY	250	GATGAACCTAATGAATCAATCAGTCTTGGACCCGCTGGAAATTTCTTCACTGAAGAG	309
Db	249	GATGAACCTAATGAATCAATCAGTCTTGGACCCGCTGGAAATTTCTTCACTGAAGAG	308
QY	310	GTGGACTCCGAAAAATTGACGAGAGGAAATCCAGATGAACACTTTCGAGAAATTCG	369

Db	309	GTGGACTCCGAAATATGACAGGAGGAAATCCAGATGAAACTTTGGAGAAATG	368	Qy	1450	ACAGTCATGATACCGTTGGCCGAGGCTTGGGACTTCCCTGGCCGAACACTGTGGACCTG	1509
Qy	370	AAGAGCCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCTGGCTTCNC	429	Db	1449	ACAGTCATGATACCGTTGGCCGAGGCTTGGGACTTCCCTGGCCGAACACTGTGGACCTG	1508
Db	369	AAGAGCCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCTGGCTTCNC	428	Qy	1510	GGGCTGACAGGCAACCATGAGTTGTGCACCCAGTCTTGGGACAGTGCACCAAGTTT	1569
Qy	430	AACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGGTCCATCACTGTACC	489	Db	1509	GGGCTGACAGGCAACCATGAGTTGTGCACCCAGTCTTGGGACAGTGCACCAAGTTT	1568
Db	429	AACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGGTCCATCACTGTACC	488	Qy	1570	GAGGAGAACACTACTCTTGGCCGAGCGTGGAGACACTGTGTCTCCCTTTGGCAAG	1629
Qy	490	CTGGCAGCGCACCAGGCTATTGGCTTCAAGGGGATCATCTTGGCTGGCACTGAGAGCAG	549	Db	1569	GAGGAGAACACTACTCTTGGCCGAGCGTGGAGACACTGTGTCTCCCTTTGGCAAG	1628
Db	489	CTGGCAGCGCACCAGGCTATTGGCTTCAAGGGGATCATCTTGGCTGGCACTGAGAGCAG	548	Qy	1630	ACCATCATGAGAGGAGCAGCTGTGTAAGCGGGTGGCCACACTCTCATCAACCTGTAT	1689
Qy	550	AAAGCCAAATPACTTGGCTTAACCTGAGCGTTCGGGGAGCACAATTGCAGCCTTCTGCCTCAG	609	Db	1629	ACCATCATGAGAGGAGCAGCTGTGTAAGCGGGTGGCCACACTCTCATCAACCTGTAT	1688
Db	549	AAAGCCAAATPACTTGGCTTAACCTGAGCGTTCGGGGAGCACAATTGCAGCCTTCTGCCTCAG	608	Qy	1690	GGCATGAGCGCGTGTCTGTCGGCGGCGAGCGCTCCATCCGCAATGGGCTCCGCAACAC	1749
Qy	610	GAGCCAGCAGTGGGAGCGATGAGGCTCAATCCGAGGAGAGCCACACTAAGTGAAGAC	669	Db	1689	GGCATGAGCGCGTGTCTGTCGGCGGCGAGCGCTCCATCCGCAATGGGCTCCGCAACAC	1748
Db	609	GAGCCAGCAGTGGGAGCGATGAGGCTCAATCCGAGGAGAGCCACACTAAGTGAAGAC	668	Qy	1750	GACCAAGAGGTTCTTGGCCAAACACTTCTTGGTGGAGCTTACTTGCAGAACTCTCTC	1809
Qy	670	AAGAGCAGTACACTTCAATGGCTTCAAGGCTTCAAGGCTTCAATCCGAGGAGTGGCCAT	729	Db	1749	GACCAAGAGGTTCTTGGCCAAACACTTCTTGGTGGAGCTTACTTGCAGAACTCTCTC	1808
Db	669	AAGAGCAGTACACTTCAATGGCTTCAAGGCTTCAATCCGAGGAGTGGCCAT	728	Qy	1810	AGCCTCTCTCAGCTGGCAAGTATGCTCCAGAAACCTTAGATGAGCAGATTAAGAAAGTG	1869
Qy	730	ATTTTACTGTGTTGCAAGACTGAGTGTGTTGATTTCTGATGGATCACTGAAAGACAA	789	Db	1809	AGCCTCTCTCAGCTGGCAAGTATGCTCCAGAAACCTTAGATGAGCAGATTAAGAAAGTG	1868
Db	729	ATTTTACTGTGTTGCAAGACTGAGTGTGTTGATTTCTGATGGATCACTGAAAGACAA	788	Qy	1870	TCCCAAGCAGTCTTGAAGGAGCGCTATATCTGTGCGCCACCTCTGGACAGCAGATGC	1929
Qy	790	ATCAGCATTCATPAGTAGAAGAGACTTGGTGGAGTCACTAATGGGAAACCCGAGAT	849	Db	1869	TCCCAAGCAGTCTTGAAGGAGCGCTATATCTGTGCGCCACCTCTGGACAGCAGATGC	1928
Db	789	ATCAGCATTCATPAGTAGAAGAGACTTGGTGGAGTCACTAATGGGAAACCCGAGAT	848	Qy	1930	TGAGCAGGAGCAGTCTCCCTGTACCGCCGCGCTTACCCATGGCCGCTTGTGTGAT	1989
Qy	850	AAATAGCATTTCGGGCTCCACACTTGTGAAGTCCATTTTGAAGACCAAGATACCT	909	Db	1929	TGAGCAGGAGCAGTCTCCCTGTACCGCCGCGCTTACCCATGGCCGCTTGTGTGAT	1988
Db	849	AAATAGCATTTCGGGCTCCACACTTGTGAAGTCCATTTTGAAGACCAAGATACCT	908	Qy	1990	GACTGTACTCTTTTTCAGAAAGGTTGGGATTTATCACAGGTTAAGCTTTTGTTCCTC	2049
Qy	910	GTGGAACATCTTGGAGAGTCCGAGATGGGTTTAAAGTGGCCATGAACATCCCTCAC	969	Db	1989	GACTGTACTCTTTTTCAGAAAGGTTGGGATTTATCACAGGTTAAGCTTTTGTTCCTC	2048
Db	909	GTGGAACATCTTGGAGAGTCCGAGATGGGTTTAAAGTGGCCATGAACATCCCTCAC	968	Qy	2050	GTCTGCACCTGAAGGTTCTCGCCTGGCCTGGAGAGCCTTCCAGGTTTGTGACCTGCA	2109
Qy	970	ACCGCCGGTTCAGCATGGCAGCGTCTGGCTGGCTTCAAGAGATTGATGAATG	1029	Db	2049	GTCTGCACCTGAAGGTTCTCGCCTGGCCTGGAGAGCCTTCCAGGTTTGTGACCTGCA	2108
Db	969	ACCGCCGGTTCAGCATGGCAGCGTCTGGCTGGCTTCAAGAGATTGATGAATG	1028	Qy	2110	GGCAGTGTCTCTTAACAGGACCATCACAGCTTCTGAACCTGAGCCGAGAGAGAGATGGA	2169
Qy	1030	ACTGCTGAGTACGCTGCACAGGAACAGTTTAAACAGAGGCTCAGTGAATTTGGATTG	1089	Db	2109	GGCAGTGTCTCTTAACAGGACCATCACAGCTTCTGAACCTGAGCCGAGAGAGAGATGGA	2168
Db	1029	ACTGCTGAGTACGCTGCACAGGAACAGTTTAAACAGAGGCTCAGTGAATTTGGATTG	1088	Qy	2170	ATTGCTGACCCCTGGAACTGGCGGTTATTTCTGGTCAATTGAGGAGACCATAGTGAAC	2229
Qy	1090	ATTGAGGAAATTTGCACTGATGGCTCAGAGGCTTACGTCAATGGAGATGACCTAC	1149	Db	2169	ATTGCTGACCCCTGGAACTGGCGGTTATTTCTGGTCAATTGAGGAGACCATAGTGAAC	2228
Db	1089	ATTGAGGAAATTTGCACTGATGGCTCAGAGGCTTACGTCAATGGAGATGACCTAC	1148	Qy	2230	TGGGGCTTATGCTGCTCCCTCCAGGTTGTGAGTGGTGGGACCTGTGTGAGGTGGA	2289
Qy	1150	CTCAGAGGAGGATCTGACCAACCTGGCTTCCCGACTGCTCCATCGAGGAGCCCATG	1209	Db	2229	TGGGGCTTATGCTGCTCCCTCCAGGTTGTGAGTGGTGGGACCTGTGTGAGGTGGA	2288
Db	1149	CTCAGAGGAGGATCTGACCAACCTGGCTTCCCGACTGCTCCATCGAGGAGCCCATG	1208	Qy	2290	TAGCCATTTCTGCTTCAACACCATCTCTTAAGAAACAGCTTCAAGAGCTGTGTGCTG	2349
Qy	1210	GTGAAGGTTTCACTGCTGAGCGCCGTGGCAGTGTGTGAGTGGCCCTGCAGATCCTC	1269	Db	2289	TAGCCATTTCTGCTTCAACACCATCTCTTAAGAAACAGCTTCAAGAGCTGTGTGCTG	2348
Db	1209	GTGAAGGTTTCACTGCTGAGCGCCGTGGCAGTGTGTGAGTGGCCCTGCAGATCCTC	1268	Qy	2350	ATTCAATTTAACTAGAGCAGGACCTTAAACATGTACCGAGAACCACTTTAAACAAAGA	2409
Qy	1270	GGGGCTTGGCTTACACAGGAGCTATCCGTACGAGGCGATCTCGTGAACCCGCAATC	1329	Db	2349	ATTCAATTTAACTAGAGCAGGACCTTAAACATGTACCGAGAACCACTTTAAACAAAGA	2408
Db	1269	GGGGCTTGGCTTACACAGGAGCTATCCGTACGAGGCGATCTCGTGAACCCGCAATC	1328	Qy	2410	ATATAAATGTCACACTCTGTACTGTTAAAAAATAA 2449	
Qy	1330	CTCCTCATCTTCAGGGAACCAATGAGATTCTCCGATGTATACCTCGCCCTGACGGGTG	1389	Db	2409	ATATAAATGTCCCAATCAAAAAAATAAATAAATAA 2448	
Db	1329	CTCCTCATCTTCAGGGAACCAATGAGATTCTCCGATGTATACCTCGCCCTGACGGGTG	1388	Qy			
Qy	1390	CAGCATCGCGCGCATCTGACTACAGGATCCATGAGCTTAAACAGCAGCAAGTGAGC	1449	Db			
Db	1389	CAGCATCGCGCGCATCTGACTACAGGATCCATGAGCTTAAACAGCAGCAAGTGAGC	1448				

RESULT 4
AF327351
LOCUS

AF327351 2411 bp mRNA llinear PRI 30-DEC-2002

DEFINITION

Homo sapiens very-long-chain acyl-CoA dehydrogenase VLCAD mRNA, complete cds.

ACCESSION

AF327351

VERSION

AF327351.1 GI:18028282

KEYWORDS

Source

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 2411)
Zhang, J., Zhang, W., Zou, D., Chen, G., Wan, T., Zhang, M. and Cao, X.
Cloning and functional characterization of ACAD-9, a novel member
of human acyl-CoA dehydrogenase family
Biochem. Biophys. Res. Commun. 297 (4), 1033-1042 (2002)

AUTHORS

Zhang, J., Zhang, W., Zou, D., Chen, G., Wan, T., Zhang, M. and Cao, X.

TITLE

Direct Submission

JOURNAL

Submitted (05-DEC-2000)

MEDLINE

Department of Immunology, Second Military
Medical University & Shanghai Brilliance Biotechnology Institute,
800 Xiangyin Rd., Shanghai 200433, P.R. China

PUBMED

Location/Qualifiers

REFERENCE

1..2411

AUTHORS

/organism="Homo sapiens"

JOURNAL

/mol_type="mRNA"

FEATURES

/db_xref="taxon:9606"

source

33..1898

CDS

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

protein_id="AAL56011.1"

/db_xref="GI:18028283"

translation="MSGGGLFLRTTAAARACGLVSTANRLRLTSPVRAFAKELF

LGKIRKKEVPEPEVSDLEINOFPLGVEKFFTEEDVSRKIDDEGKIPDETLEKLG

SLGLFLQVPEYEGYGLFNTWYSLRGETISMDGSITVTLLAQAIGLKLGLAGTTE

QAKYLPKLASGEHIAFLCTEPASGDAASRATLSKDKHILNGSKVWITNGG

LANITVFATKTEVSDGSKDITAFIVERDFGVGTNGKPEDKILGRSNTCEVHFE

NTKIVENILGEVGGFKVMNSGRFSVAGLLKRLIEMTARYACTRQFNK

RUSEGLIOEKFAKMAQKAYNMESYITAGMLDQPPDCSIEAMVKVFSSEAOW

CVEYLIQIGGLGYTRDYPERILRDLRIILFETNEILRWYIALTGLQHAGRLIT

RHELKQAKVSTYMDTVGRRLRSLGRTVDLGTNGHGVHPSLADSNKFEENTYCF

GRVETLLRFKFTLMEQLVLRKRVANILINLYGMTAVLSRPSRSIRIGLNKHDEVL

LANTCEVAYLQNLSELSOLDKYAPENLDEIKKVSQILKEKRAYICAHPLDRTC"

612 a 604 c 609 g 526 t

BASE COUNT

612 a 604 c 609 g 526 t

ORIGIN

612 a 604 c 609 g 526 t

Query Match

Best Local Similarity 97.5%; Score 2391.8; DB 9; Length 2411;

Matches 2399; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

35 ACTGAGGCTGAGGCTGGGGAACATCGGCGACATGAGCGGCTGCGGGCTCTTCCTGGCA 94

1 ACTGAGGCTGAGGCTGGGGAACATCGGCGACATGAGCGGCTGCGGGCTCTTCCTGGCA 60

95 CCACGGCTGCGGCTGCTGCCCTCGCGGGCTCTGGTCTCTACCGCAACCGCGGCTAC 154

61 CCACGGCTGCGGCTGCTGCCCTCGCGGGCTCTGGTCTCTACCGCAACCGCGGCTAC 120

155 TCGCGCACCGCCGCTGACAGCTTTCGCCCAAGAGCTTTTCTAGGCAAAATCAAGA 214

131 TCGCGCACCGCCGCTGACAGCTTTCGCCCAAGAGCTTTTCTAGGCAAAATCAAGA 180

215 AGAAGAAGTTTTCCTCCATTTCCAGAGTTAGCCCAAGATGAACITTAATGAATCAATCAGT 274

181 AGAAGAAGTTTTCCTCCATTTCCAGAGTTAGCCCAAGATGAACITTAATGAATCAATCAGT 240

275 TCTTGGGACCGCTGGAAAAATCTTCTACTCAAGAGTGGACTCCCGAAAAATGACACAGG 334

241 TCTTGGGACCGCTGGAAAAATCTTCTACTCAAGAGTGGACTCCCGAAAAATGACACAGG 300

335 AAGGAAAAATCCAGATGAACITTTGAGAAAAATGAGAGCCCTAGGCGCTTTTGGGCTGC 394

301 AAGGAAAAATCCAGATGAACITTTGAGAAAAATGAGAGCCCTAGGCGCTTTTGGGCTGC 360

395 AAGTCCCAAGAATAATGCTGGCTGGCTTCTCCAACACCATGTTACTCAAGACTAGGGG 454
|||||
361 AAGTCCCAAGAATAATGCTGGCTGGCTTCTCCAACACCATGTTACTCAAGACTAGGGG 420
|||||
455 AGATCATCAGCATGGATGGTCCATCACTGTGACCTGGCAGCGCACAGGCTATTGGCC 514
421 AGATCATCAGCATGGATGGTCCATCACTGTGACCTGGCAGCGCACAGGCTATTGGCC 480
|||||
515 TCAAGGGATCATCTTGGCTGGCACTGAGGACAGAAAGCAAAATGCTTAACTGG 574
481 TCAAGGGATCATCTTGGCTGGCACTGAGGACAGAAAGCAAAATGCTTAACTGG 540
|||||
575 CTTCCGGGAGCACATTCAGCTTCTCCCTCAGGAGCGACGAGCTGGGAGCCATG 634
541 CTTCCGGGAGCACATTCAGCTTCTCCCTCAGGAGCGACGAGCTGGGAGCCATG 600
|||||
635 CTTCAATCCGGAGCAGCCACACTAAGTGAAGCAAGAACACTACATCTCAATGGCT 694
601 CTTCAATCCGGAGCAGCCACACTAAGTGAAGCAAGAACACTACATCTCAATGGCT 660
|||||
695 CCAAGGCTGCTGATTAATGAGGACTGGCCCAATATTTTACTGTGTTGCAAAAGCTG 754
661 CCAAGGCTGCTGATTAATGAGGACTGGCCCAATATTTTACTGTGTTGCAAAAGCTG 720
|||||
755 AGGCTGCTGATTCATGATGATGATGAAAGCAAAATCAGAGCTTGGGGCTTCCAAACA 814
721 AGGCTGCTGATTCATGATGATGATGAAAGCAAAATCAGAGCTTGGGGCTTCCAAACA 780
|||||
815 ACTTTGGTGGAGTCACTAATGGGAAACCCGAGAAATTTAGGCTTTCGGGGCTTCCAAACA 874
781 ACTTTGGTGGAGTCACTAATGGGAAACCCGAGAAATTTAGGCTTTCGGGGCTTCCAAACA 840
|||||
875 CTTGTAAGTCCATTTTGAAGCAACCAAGATACCTGTGGAAGCAATCTTGGAGAGTGC 934
841 CTTGTAAGTCCATTTTGAAGCAACCAAGATACCTGTGGAAGCAATCTTGGAGAGTGC 900
|||||
935 GAGATGGGTTAAGTGGCCATGAACATCTCAACAGCGCCGGTTCAGCATGGGAGCG 994
901 GAGATGGGTTAAGTGGCCATGAACATCTCAACAGCGCCGGTTCAGCATGGGAGCG 960
|||||
995 TCGTGGCTGGGCTGCTCAAGAGATTTGATTAAGTGAATGACTGCTGAGTACGCTGCACAGGA 1054
961 TCGTGGCTGGGCTGCTCAAGAGATTTGATTAAGTGAATGACTGCTGAGTACGCTGCACAGGA 1020
|||||
1055 AACAGTTTAAACAGAGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCACGTATGG 1114
1021 AACAGTTTAAACAGAGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCACGTATGG 1080
|||||
1115 CTGAGAGGCTTACGTCATGGAGATGACCTACCTACCTACAGAGGATGCTGGAGCCAAAC 1174
1081 CTGAGAGGCTTACGTCATGGAGATGACCTACCTACCTACAGAGGATGCTGGAGCCAAAC 1140
|||||
1175 CTGGCTTTCCGACTCTCTCCATCGAGGACCCATGCTGAAGGTGTTTACGCTCCGAGGCGG 1234
1141 CTGGCTTTCCGACTCTCTCCATCGAGGACCCATGCTGAAGGTGTTTACGCTCCGAGGCGG 1200
|||||
1235 CTTGGCAGTGTGAGTGGAGCGCTGCAGATCCTCGGGGCTTGGGGCTTACAGAGGACT 1294
1201 CTTGGCAGTGTGAGTGGAGCGCTGCAGATCCTCGGGGCTTGGGGCTTACAGAGGACT 1260
|||||
1295 ATCCGTACGAGCGCATCTGCGTGACACCCGATCCTCTCATCTTCCAGGGAACCAATG 1354
1261 ATCCGTACGAGCGCATCTGCGTGACACCCGATCCTCTCATCTTCCAGGGAACCAATG 1320
|||||
1355 AGATTCTCGGATGTACATGCGCCCTGAGCGGTCTGACAGCTGCGGCGGCTTCTGACTA 1414
1321 AGATTCTCGGATGTACATGCGCCCTGAGCGGTCTGACAGCTGCGGCGGCTTCTGACTA 1380
|||||
1415 CCAGGATCCATGAGCTTAAACAGGCAAGTGTGAGCAGTGTGATGATCCGTTGGCGCGGA 1474
1381 CCAGGATCCATGAGCTTAAACAGGCAAGTGTGAGCAGTGTGATGATCCGTTGGCGCGGA 1440

AUTHORS	Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wagatsuma, M., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, N., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.
TITLE	NEDO human cDNA sequencing project
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 2399)
AUTHORS	Isogai, T. and Otsuki, T.
TITLE	Direct Submission
JOURNAL	Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 252-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
FEATURES	Location/Qualifiers
source	1..2399
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="Y79AA1001048"
	/cell_line="Y79"
	/cell_type="retinoblastoma"
	/clone_lib="Y79AA1"
	/note="cloning vector: pME18SFL3"
	47..1912
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="BAB14775.1"
	/db_xref="GI:10436258"
CDS	translation="MSGGLFLRTAAARCGLVSTANRLRLTSPVRAFAKELF LKIKKKVFPEVPSQDELNEINQFLPVEKFTPEVDSKKIDQEKIDPELEKLSIGLEFLQVPEYGGISFNTMYSLRGILISMDGSIITVLAHQALGKILLIAGTE QKALYPLKLSAGEHSIAFLPEPASGSDAISRATLSDEDKHYILUNGSKVITNGG LANITFVAKTEVDSDGVRDKITAFIVERDFGVNGKPEDKLGRSGNTEVHEE NTKIPVENLGEVGFQKVAYNILNSGFSVMSVAGLLKRLIEMTAYACTRKQFNK RLSEFLGLKGFALMAQKAYVMSMTYLTAGMLDQPGFPDCSIEAAMVKVFSSEAAWQ CVSEALQILGGLYTVDYPERILDRILIFECTNEILRMVIALTGLQHGRLIT RTHETKOAKSTVMDTGVRLRDSIGRTVDLGTGNHGVHPHSLADSANKFEENTYCF GRIVETLLLRGKILMERQVLVLRVANILINLYGWTAVLSRPSRSIRIGLRNHDHVEL LANTCEVAYLQNLFLSOLDKRYAPENLDEQIKKVSQOILEKRAYICAHPLDRFC"
BASE COUNT	597 a 606 c 672 g 524 t
ORIGIN	
Query Match	97.2%; Score 2382.2; DB 9; Length 2399;
Best Local Similarity	99.8%; Pred. No. 0;
Matches 2396; Conservative	0; Mismatches 3; Indels 2; Gaps 1;
QY	21 GCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGAACATCGGSCAGCATGAGCGGCTGGG 80
Db	1 GCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGAACATCGGSCAGCATGAGCGGCTGGG 60
QY	81 GCTCTTCCTGCCACACGCGTGGCGTCGTCGCTGCCGGGCTGCTGCTACCGC 140
Db	61 GCTCTTCCTGCCACACGCGTGGCGTCGTCGCTGCCGGGCTGCTGCTACCGC 120
QY	141 GAACCGCGGCTACTGCGCACCGCGCTGTACGAGCTTTCGCCAAGAGCTTTTCC 200
Db	121 GAACCGCGGCTACTGCGCACCGCGCTGTACGAGCTTTCGCCAAGAGCTTTTCC 180
QY	201 AGCCAAATTCAGAGAAGAAGATTTCCTATTTCCAGAGTTACCCAGATGACTTAA 260
Db	181 AGCCAAATTCAGAGAAGAAGATTTCCTATTTCCAGAGTTACCCAGATGACTTAA 240
QY	261 TGAATCAATCAGTTCTTGGGACCGCTGGAAAAATTTCTACTGAAGAGGTGGACTCCG 320
Db	1321 CGAGGGAACCAATGAGATTCTCGGATGTACATCGCCCTGACGGGTCTGCAGCATCCCG 1380
Db	1341 CGAGGGAACCAATGAGATTCTCGGATGTACATCGCCCTGACGGGTCTGCAGCATCCCG 1400
Db	1261 CTACACAAGGAGCTATCCGTACGAGCGCATATTTCGTGACACCCGCATCTCTCATCTT 1320
QY	1281 CTACACAAGGAGCTATCCGTACGAGCGCATATTTCGTGACACCCGCATCTCTCATCTT 1340
Db	1201 CAGCTCGAGCGCGCTGGCAGTGTGTGAGTGAGGCGTGCAGATCTCTCGGGGCTTGG 1260
QY	1161 GATCTGGACCAACCTGGCTTTCCGACTGCTCCATCGAGGCGACCATGGTGAAGTGT 1220
Db	1141 GATGCTGGACCAACCTGGCTTTCCGACTGCTCCATCGAGGCGACCATGGTGAAGTGT 1200
QY	1041 CGCCTGCACAAGGAAACAGTTTAAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAA 1100
Db	1021 CGCCTGCACAAGGAAACAGTTTAAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAA 1080
QY	1101 ATTTGCACTGATCGCTCAGAAGGCTTACGTCATGGAGAGTATGACCTACTCACAGCAG 1160
Db	1081 ATTTGCACTGATCGCTCAGAAGGCTTACGTCATGGAGAGTATGACCTACTCACAGCAG 1140
QY	901 CTTGGAGAGGTCGGAGATGGGTTTAAAGTGGCCATGAACATCTCTCAACAGCGCGGTT 960
Db	981 CAGCATGGGAGGTCGCTGGCTGGCTGCTCAAGAGATTGATTGAATGACTGCTGAGTA 1040
QY	841 TCGGGGCTCCAACACATTTGTGAAGTCCATTTTGAACACACCAAGATACCTGTGAAAACAT 900
Db	861 TCAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAGATAAATTAGGCAT 880
QY	781 CATCTCAATGGCTCCAAGGCTGGATTACTAATGGAGGACTGGCCAAATATTTTACTGT 740
Db	721 GTTTGCAAGAGACTGAGTCTGTTGATTCTCTGATGGATCAGTGAAGACAAATCACAGCAT 800
QY	661 CATCTCAATGGCTCCAAGGCTGGATTACTAATGGAGGACTGGCCAAATATTTTACTGT 720
Db	601 TGGGAGCGGATGCAGCCTCAATCCGAGAGCAGCCACTAGTAGAAGACAGAGCAGCTA 660
QY	561 CTTGCCCTAACTGGGCTCCGGGAGGACACATTCAGCCTTCTGCTCAGGAGCCAGCCAG 620
Db	541 CTTGCCCTAACTGGGCTCCGGGAGGACACATTCAGCCTTCTGCTCAGGAGCCAGCCAG 600
QY	481 CCAGGCTATTGGCCTCAAGGGATCATCTTGGCTGGCCTAGAGGAGCAAGAACCAATA 540
Db	461 CCAGGCTATTGGCCTCAAGGGATCATCTTGGCTGGCCTAGAGGAGCAAGAACCAATA 520
QY	421 CTCACGACTAGGGGAGATCATCAGATGGATGGTCCATCACTGTGACCTTGGCAGCGCA 480
Db	401 CTCACGACTAGGGGAGATCATCAGATGGATGGTCCATCACTGTGACCTTGGCAGCGCA 460
QY	361 GCTTTTGGGCTGCAAGTCCCAAGAAATATGGTGGCTGGGCTTCTCCACACCATGTA 420
Db	341 GCTTTTGGGCTGCAAGTCCCAAGAAATATGGTGGCTGGGCTTCTCCACACCATGTA 400
QY	321 AAAAAATTGACAGAGGAGGAAAAATCCAGATGAAACTTTGGAGAAATTTGAAGAGCCTAGG 380
Db	301 AAAAAATTGACAGAGGAGGAAAAATCCAGATGAAACTTTGGAGAAATTTGAAGAGCCTAGG 360
QY	241 TGAATCAATCAGTTCTTGGGAGCCCGTGGAAAAATTTCTTCACTGAAGAGGTGGACTCCG 300

QY 1401 CGGCATCTCTGACTACCGAGATCCATGAGCTTAAACAGCCAAAGTGGACACAGTCATGGA 1460
 DB 1381 CGGCATCTCTGACTACCGAGATCCATGAGCTTAAACAGCCAAAGTGGACACAGTCATGGA 1440
 QY 1461 TACCGTTGGCGGAGGCTTCGGGACTCCCTGGCGGAACTGTGGACCTGGGCTGACAGG 1520
 DB 1441 TACCGTTGGCGGAGGCTTCGGGACTCCCTGGCGGAACTGTGGACCTGGGCTGACAGG 1500
 QY 1521 CAACCATGGAGCTTGTGACCCAGCTCTTGGGACAGTGCCAAAGTTTGGAGAGAACAC 1580
 DB 1501 CAACCATGGAGCTTGTGACCCAGCTCTTGGGACAGTGCCAAAGTTTGGAGAGAACAC 1560
 QY 1581 CTACTGCTTCGGCGGAGGCTTCGGGACTGCTGCTCCGCTTTGGCAAGACCATCATGGA 1640
 DB 1561 CTACTGCTTCGGCGGAGGCTTCGGGACTGCTGCTCCGCTTTGGCAAGACCATCATGGA 1620
 QY 1641 GGACACGCTGTGACTGAAGCGGCTTCGGGAGGCTTCATCAACCTGTATGGCATGACGGC 1700
 DB 1621 GGACACGCTGTGACTGAAGCGGCTTCGGGAGGCTTCATCAACCTGTATGGCATGACGGC 1680
 QY 1701 CGTCTCTCGCGGAGGCTTCGCTCCATCCGCTTCGCTCCGCTCCGCAACGACGACAGGT 1760
 DB 1681 CGTCTCTCGCGGAGGCTTCGCTCCATCCGCTTCGCTCCGCTCCGCAACGACGACAGGT 1740
 QY 1761 TCTCTTGGCAACACCTTCTGCTGGAAGCTTACTTGCAGAACTCTTTCAGGCTCTCTCA 1820
 DB 1741 TCTCTTGGCAACACCTTCTGCTGGAAGCTTACTTGCAGAACTCTTTCAGGCTCTCTCA 1800
 QY 1821 GCTGGCAAGTATGCTCCAGAAACCTAGATGAGCAGATTAAGAAGTGCTCCAGCAGAT 1880
 DB 1801 GCTGGCAAGTATGCTCCAGAAACCTAGATGAGCAGATTAAGAAGTGCTCCAGCAGAT 1860
 QY 1881 CCTTGAGAGCGAGGCTATATCTGTCGCCACCTCTGGACAGGACATGCTGAGGACGAGG 1940
 DB 1861 CCTTGAGAGCGAGGCTATATCTGTCGCCACCTCTGGACAGGACATGCTGAGGACGAGG 1920
 QY 1941 ACAGTCTCCCTCTACCGCCGCGCTTACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2000
 DB 1921 ACAGTCTCCCTCTACCGCCGCGCTTACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
 QY 2001 TTTTTCAGAGGCTGTGGGATTAACAGGTTTAAAGCTTTTGTTCCTCCGCTGCTGCTGCTG 2060
 DB 1981 TTTTTCAGAGGCTGTGGGATTAACAGGTTTAAAGCTTTTGTTCCTCCGCTGCTGCTGCTG 2040
 QY 2061 AAGGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2120
 DB 2041 AAGGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
 QY 2121 CTAAAGGACCATCAAGCTTCTGAACTGAGCGGAGAGAGAGATGAATGCTGCTGCTGCTG 2180
 DB 2101 CTAAAGGACCATCAAGCTTCTGAACTGAGCGGAGAGAGAGATGAATGCTGCTGCTGCTG 2160
 QY 2181 CTGGAAGTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2240
 DB 2161 CTGGAAGTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
 QY 2241 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2300
 DB 2221 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
 QY 2301 GCTCAACACACATCTCTAAGAAACAGCTTGAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 2360
 DB 2279 GCTCAACACACATCTCTAAGAAACAGCTTGAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
 QY 2361 CTAGAAGCAGAGGCTTAAACATGTACCAGGACCATTTAACAAGAAATATAAATGT 2420
 DB 2339 CTAGAAGCAGAGGCTTAAACATGTACCAGGACCATTTAACAAGAAATATAAATGT 2400
 QY 2421 C 2421
 DB 2399 C 2399

RESULT 7
 BD157796
 LOCUS
 DEFINITION
 BD157796
 ACCESSION
 BD157796.1 GI:27863554
 VERSION
 JP 2002191363-A/12639
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2400)
 Ota.T., Isogai.T., Nishikawa.T., Hayashi.K., Saito.K., Yamamoto.J.,
 Ishii.S., Sugiyama.T., Wakamatsu.A., Nagai.K. and Otsuki.T.,
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 12639 09-JUL-2002;
 HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002191363-A/12639
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
 PI SAITO
 PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
 PI KEIICHI NAGAI,TETSUJI OTSUKI
 PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
 10,
 PC C12P21/02,C12O1/68/C12P21/08,C06F17/30,C12N15/00,C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH key

FT CDS Location/Qualifiers
 (498)..(1991).
 1..2400
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 570 a 629 c 685 g 516 t
 ORIGIN

Query Match 86.3%; Score 2115.4; DB 6; Length 2400;
 Best Local Similarity 94.2%; Pred. No. 0;
 Matches 2262; Conservative 0; Mismatches 56; Indels 83; Gaps 3;
 QY 21 GCTAAGAGGGAGAGCTGAGGCTGGGGAACATCGGCACATGAGCGGCTGGCG 80
 DB 1 GCTAAGAGGGAGAGCTGAGGCTGGGGAACATCGGCACATGAGCGGCTGGCG 60
 QY 81 GCTCTTCTGCGCACCACCGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCT 140
 DB 61 GCTCTTCTGCGCACCACCGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCT 120
 QY 141 GAACCGGCGGCTACTGCGCACCACCGGCTGCTGAGGCTTTGCGCAAGAGCTTTTCT 200
 DB 121 GAACCGGCGGCTACTGCGCACCACCGGCTGCTGAGGCTTTGCGCAAGAGCTTTTCT 180
 QY 201 AGCAAAATCAAGAAGAAA----- 219
 DB 181 AGCAAAATCAAGAAGGTAACCGGAGGCTTGGGGAACCTTGTCTGCTGCTGCTGCT 240
 QY 220 -----GAAGTTTTCCTTTTCCAGAGTTTTCAGAGTTTTCAGAGTTTTCAGAGTT 252
 DB 241 TTCACCCCTCAGCTGCAAGACTGGTGTGAACTTTTGTGAGATTTCCCAACCTGCCAGAGA 300
 QY 253 GAACCT-----TAATGAATCAATCAGTTCTTGGGACCGCTGGAAATTTCT 298
 DB 301 GATACACCCCTGGGCGGAGGCTGTAACTCGGGATTTCTGAGTTTCCAGGAAACCTT 360
 QY 299 TCACCTGAAGAGGTGGACTCCCGAAAAATTGACAGGAGGAGAAATCCAGATGAACCTT 358
 DB 361 CCAGAGAGAGGTGGACTCCCGAAAAATTGACAGGAGGAGAAATCCAGATGAACCTT 420

QY 359 TGGAGAAATGAAGAGCTAGGGCTTTTGGCTGCAAGTCCAGAGAAATATGGTGGCC 418
Db 421 TGGAGAAATGAAGAGCTAGGGCTTTTGGCTGCAAGTCCAGAGAAATATGGTGGCC 480
QY 419 TGGCTTCTCAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGGTCCA 478
Db 481 TGGCTTCTCAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGGTCCA 540
QY 479 TCACTGTGACCTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCA 538
Db 541 TCACTGTGACCTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCA 600
QY 539 CTGAGGAGCAAGAGCCAAATACTTGGCTAAACTGGCGTCCGGGGAGCACATTGAGCCT 598
Db 601 CTGAGGAGCAAGAGCCAAATACTTGGCTAAACTGGCGTCCGGGGAGCACATTGAGCCT 660
QY 599 TCTGCCTCAGCGGAGCCAGCTAGTGGAGCGATCAGGCTCAATCCGGAGCAGAGCCACAC 658
Db 661 TCTGCCTCAGCGGAGCCAGCTAGTGGAGCGATCAGGCTCAATCCGGAGCAGAGCCACAC 720
QY 659 TAAGTGAAGACAAGAGCACTACATCTCAATGGCTCCAAGGCTCGGATTAATGGAG 718
Db 721 TAAGTGAAGACAAGAGCACTACATCTCAATGGCTCCAAGGCTCGGATTAATGGAG 780
QY 719 GACTGGCCAAATATTTTACTGTGTGCAAGAGACTGAGTGTGATCTGATGGATCAG 778
Db 781 GACTGGCCAAATATTTTACTGTGTGCAAGAGACTGAGTGTGATCTGATGGATCAG 840
QY 779 TCAAGAGACAATAACACGATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGA 838
Db 841 TGAAGAGACAATAACACGATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGA 900
QY 839 AACCCGAGATAAATAGGCAATTCGGGCTCCACACTGTGGAAGTCCATTTTGAAGACA 898
Db 901 AACCCGAGATAAATAGGCAATTCGGGCTCCACACTGTGGAAGTCCATTTTGAAGACA 960
QY 899 CCAAGTACCTGTGGAAGAACATCTTGGAGAGTGGGATTTAAGTGGCCATGA 958
Db 961 CCAAGTACCTGTGGAAGAACATCTTGGAGAGTGGGATTTAAGTGGCCATGA 1020
QY 959 ACATCCTCAACAGCGCGGTTTACAGTGGGAGGCGTGGTGGGCTGCTCAAGAGAT 1018
Db 1021 ACATCCTCAACAGCGCGGTTTACAGTGGGAGGCGTGGTGGGCTGCTCAAGAGAT 1080
QY 1019 TCATTGAATGACTGCTGAGTACGCTGCACAAGAGAAACAGTTTAAACAGAGGCTCAGTG 1078
Db 1081 TGAATGAATGACTGCTGAGTACGCTGCACAAGAGAAACAGTTTAAACAGAGGCTCAGTG 1140
QY 1079 AATTTGGATTGATCAGGAGAAATTTGCACTGATGGCTCAGAAGCCTTACGTCATGGAGA 1138
Db 1141 AATTTGGATTGATCAGGAGAAATTTGCACTGATGGCTCAGAAGCCTTACGTCATGGAGA 1200
QY 1139 GTATGACCTACCTCAGCAGGAGTGTGGACCAACCTGGCTTTCCGCACTGCTCCATCG 1198
Db 1201 GTATGACCTACCTCAGCAGGAGTGTGGACCAACCTGGCTTTCCGCACTGCTCCATCG 1260
QY 1199 AGGAGCCATGTTGAGTGTGTTACGCTCCGAGGCGGCTGGCAGTGTGTGAGTGAAGCGC 1258
Db 1261 AGGAGCCATGTTGAGTGTGTTACGCTCCGAGGCGGCTGGCAGTGTGTGAGTGAAGCGC 1320
QY 1259 TGCAGATCTCGGGGCTTGGGCTACACAAGGAGTATCCGTCACGAGCGCATACGCGTG 1318
Db 1321 TGCAGATCTCGGGGCTTGGGCTACACAAGGAGTATCCGTCACGAGCGCATACGCGTG 1380
QY 1319 ACACCCGCATCTCTCTCATCTTCGAGGGAACCAATCAGATTTCCCGGATGTACATCGGCC 1378
Db 1381 ACACCCGCATCTCTCTCATCTTCGAGGGAACCAATCAGATTTCCCGGATGTACATCGGCC 1440
QY 1379 TGACGGGTGTGAGCATGCGGCGGCTCCTGACTACAGGATCCATGAGCTTAAACAGG 1438
Db 1441 TGACGGGTGTGAGCATGCGGCGGCTCCTGACTACAGGATCCATGAGCTTAAACAGG 1500
QY 1439 CCAAGTGAAGACAGTCAATGGATACCGTTGGCGGAGGCTTCGGGACTCCCTTGGGCCGAA 1498

Db 1501 CCAAGTGAAGACAGTCAATGGATACCGTTGGCCGGAGGCTTCGGGACTCCCTTGGGCCAA 1560
QY 1499 CTGTGGACCTGGGCTGACAGCAACCATGGAGTTGTGCACCCCAAGTCTTGGGACAGTG 1558
Db 1561 CTGTGGACCTGGGCTGACAGCAACCATGGAGTTGTGCACCCTAGTCTTGGGACAGTG 1620
QY 1559 CCAACAAGTTTGAAGAGAACCTTACTGCTTGGCCGCGGAGCGGTGAGACACTGCTGCTCC 1618
Db 1621 CCAACAAGTTTGAAGAGAACCTTACTGCTTGGCCGCGGAGCGGTGAGACACTGCTGCTCC 1680
QY 1619 GCTTTGGCAAGACCATCATGAGGAGCAGCTGGTACTGAAGCGGTGGCCAACTCCCTCA 1678
Db 1681 GCTTTGGCAAGACCATCATGAGGAGCAGCTGGTACTGAAGCGGTGGCCAACTCCCTCA 1740
QY 1679 TCAACCTGTATGGCATGAGCGCGTGTCTCCGCGGCGCAGCGCTCATCCGCAATGGGC 1738
Db 1741 TCAACCTGTATGGCATGAGCGCGTGTCTCCGCGGCGCAGCGCTCCATCCGCAATGGGC 1800
QY 1739 TCCGCAACCAACGACGAGGTTCTTCTTGGCCCAACACCTTCTGCGTGGAAAGCTTACTTGC 1798
Db 1801 TCCGCAACCAACGACGAGGTTCTTCTTGGCCCAACACCTTCTGCGTGGAAAGCTTACTTGC 1860
QY 1799 AGAATCTTTCAGCCCTCTCTCAGCTGGACAAAGTATGCTCCAGAAAACCTTAGATGAGCAGA 1858
Db 1861 AGAATCTTTCAGCCCTCTCTCAGCTGGACAAAGTATGCTCCAGAAAACCTTAGATGAGCAGA 1920
QY 1859 TTAAGAAAGTGTCCAGCAGATCCTTGAGAAGGAGCTTATCTGTGCCACCTCTCTGG 1918
Db 1921 TTAAGAAAGTGTCCAGCAGATCCTTGAGAAGGAGCTTATCTGTGCCACCTCTCTGG 1980
QY 1919 ACAGGACATGCTGAGGAGGAGACAGTGTCCCTGCTACCGCCGCGCCCTACCCATGGCC 1978
Db 1981 ACAGGACATGCTGAGGAGGAGACAGTGTCCCTGCTACCGCCGCGCCCTACCCATGGCC 2040
QY 1979 CGTTGCTGATGACTGTACTCTTTTTCAGAGAGTGTGGGATATACAGGTTAAGCC 2038
Db 2041 CGTTGCTGATGACTGTACTCTTTTTCAGAGAGTGTGGGATATACAGGTTAAGCC 2100
QY 2039 TTTTGTCCCGCTCTCAGCTGAAGGTTGTGCGCTGGCCTGGGAGAGCTCTTCCAGGT 2098
Db 2101 TTTTGTCCCGCTCTCAGCTGAAGGTTGTGCGCTGGCCTGGGAGAGCTCTTCCAGGT 2160
QY 2099 TTTGACCTTGAGCAGTGTCTCTACAGGACCATCACAGCTTCTGAACCTGAGCGGAGA 2158
Db 2161 TTTGACCTTGAGCAGTGTCTCTACAGGACCATCACAGCTTCTGAACCTGAGCGGAGA 2220
QY 2159 GAGAGATGGAATTTGCTGACCCCTGGAACCTGGCGGTATTTCTGTCATTTGAGGAGACAC 2218
Db 2221 GAGAGATGGAATTTGCTGACCCCTGGAACCTGGCGGTATTTCT-GTCATTTGAGGAGACAC 2279
QY 2219 ATAGTGAAGACTGGGCTTATGCTGCTCCAGGCTGAGGCTGGGAGGACCTGTG 2278
Db 2280 ATAGTGAAGACTGGGCTTATGCTGCTCCAGGCTGAGGCTGGGAGGACCTGTG 2339
QY 2279 TCAGGTGTGATAGCCATTTCTGCTCAACCCACACATTTCTCTAAGAAACAGCTTGAAGCT 2338
Db 2340 TCAGGTGTGATAGCCATTTCTGCTCAACCCACACATTTCTCTAAGAAACAGCTTGAAGCT 2399
QY 2339 C 2339
Db 2400 C 2400

RESULT 8

AK022568
LOCUS Homo sapiens cDNA FLJ12506 fls, clone NT2RM2001700, weakly similar
DEFINITION to ACYL-COA DEHYDROGENASE, VERY-LONG-CHAIN SPECIFIC (EC 1.3.99.-).
ACCESSION AK022568
VERSION AK022568.1 GI:10434032
KEYWORDS oligo cloning; fls (full insert sequence).
SOURCE Homo sapiens (human)

2400 bp mRNA linear PRI 01-AUG-2002

QY	1319	ACACCGCATCTCTCTCATCTTCGAGGGAACCAATGAGATTCTCGGATGTACATCGCC	1378
DB	1381	ACACCGCATCTCTCTCATCTTCGAGGGAACCAATGAGATTCTCGGATGTACATCGCC	1440
QY	1379	TGACGGGTCTGCAGCATGCCGCCGATCTCTGACTTACAGATCATGACTTTAAACAGG	1438
DB	1441	TGACGGGTCTGCAGCATGCCGCCGATCTCTGACTTACAGATCATGACTTTAAACAGG	1500
QY	1439	CCAAAGTGACACAGTATGATACCGTTGGCCGGAGGCTTCGGGACTTCCCTGGCGCGAA	1498
DB	1501	CCAAAGTGACACAGTATGATACCGTTGGCCGGAGGCTTCGGGACTTCCCTGGCGCGAA	1560
QY	1499	CTGTGGACCTGGGGCTGCACAGCAACCATGAGTGTGTGACCCCAAGTCTTGGGGACAGTG	1558
DB	1561	CTGTGGACCTGGGGCTGCACAGCAACCATGAGTGTGTGACCCCAAGTCTTGGGGACAGTG	1620
QY	1559	CCAAACAAAGTTGAGGAGAACACTACTGCTTCGGCCGGACCGTGGAGACACTGCTGCTCC	1618
DB	1621	CCAAACAAAGTTGAGGAGAACACTACTGCTTCGGCCGGACCGTGGAGACACTGCTGCTCC	1680
QY	1619	GCCTTGGCAAGACCATCATGGAGGACAGTGGTACTGAAGGGGTGGCCAAACATCCTCA	1678
DB	1681	GCCTTGGCAAGACCATCATGGAGGACAGTGGTACTGAAGGGGTGGCCAAACATCCTCA	1740
QY	1679	TCAACCTGTATGGATGACGGCCGTCGTGTCGGCCGACCGCTCCATCCGCAATTGGGC	1738
DB	1741	TCAACCTGTATGGATGACGGCCGTCGTGTCGGCCGACCGCTCCATCCGCAATTGGGC	1800
QY	1739	TCCGCAACACGACACGAGGTCTCTTGGCCAAACACCTTCTGCGTGGGAAGCTTACTTGC	1798
DB	1801	TCCGCAACACGACACGAGGTCTCTTGGCCAAACACCTTCTGCGTGGGAAGCTTACTTGC	1860
QY	1799	AGAATCTCTCAGCCTCTCAGTGGACAAGTATGCTCCAGAAAACCTTAGATGAGCAGA	1858
DB	1861	AGAATCTCTCAGCCTCTCAGTGGACAAGTATGCTCCAGAAAACCTTAGATGAGCAGA	1920
QY	1859	TTAAGAAAGTGTCCACGACATCCTTCAGAAAGGAGCCTATATCTGTGCCACCCCTCTGG	1918
DB	1921	TTAAGAAAGTGTCCACGACATCCTTCAGAAAGGAGCCTATATCTGTGCCACCCCTCTGG	1980
QY	1919	ACAGGACATGTGAGGAGGGGACAGTGTCCCTGTACCGCCGCCCTACCACCATGGCC	1978
DB	1981	ACAGGACATGTGAGGAGGGGACAGTGTCCCTGTACCGCCGCCCTACCACCATGGCC	2040
QY	1979	CCTTCTCGATGACTGTGTACTCTTTTTCAGAAAGGTGTGGGATATACAGAGTTAAAGCC	2038
DB	2041	CCTTCTCGATGACTGTGTACTCTTTTTCAGAAAGGTGTGGGATATACAGAGTTAAAGCC	2100
QY	2039	TTTTCTTCCCGTCTGCACCTTGAAGGTTGTGCGCTTGGGAGAGCCCTTCCAGGT	2098
DB	2101	TTTTCTTCCCGTCTGCACCTTGAAGGTTGTGCGCTTGGGAGAGCCCTTCCAGGT	2160
QY	2099	TTTTGACCTTCGAGGACAGTCTCTTAACAGGACCATCACAGTCTCTGAACCTGAGCCGAGA	2158
DB	2161	TTTTGACCTTCGAGGACAGTCTCTTAACAGGACCATCACAGTCTCTGAACCTGAGCCGAGA	2220
QY	2159	GAGAGAAATGGAATGCTGACCCCTGGAACTGGCGGTATTCTGTGCTATTGAGGAGACAC	2218
DB	2221	GAGAGAAATGGAATGCTGACCCCTGGAACTGGCGGTATTCT - GTCATTGAGGAGACAC	2279
QY	2219	ATAGTGGAAACTGGGGCTTATGCTGCTGCCCTCCAGGTTGTCAGTGGTGGGGACCTGTG	2278
DB	2280	ATAGTGGAAACTGGGGCTTATGCTGCTGCCCTCCAGGTTGTCAGTGGTGGGGACCTGTG	2339
QY	2279	TCAGGTGTGGATAGGCATTTCTGCTCAACACACATTTCTTAAGAAACAGCTTGAAGACT	2338
DB	2340	TCAGGTGTGGATAGGCATTTCTGCTCAACACACATTTCTTAAGAAACAGCTTGAAGACT	2399
QY	2339	C 2339	
DB	2400	C 2400	

RESULT	9	AX394749	1863 bp	DNA	linear	PAT 18-MAY-2002
LOCUS		Sequence 3 from Patent WO0218582.				
DEFINITION		AX394749				
ACCESSION		AX394749				
VERSION		AX394749.1	G1:21065828			
KEYWORDS						
SOURCE		Homo sapiens (human)				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS		1				
TITLE		Meyers, R. and Hunter, J. J.				
JOURNAL		62112, a novel human dehydrogenase and uses thereof				
FEATURES		Patent: WO 0218582-A 3 07-MAR-2002;				
source		Millennium Pharmaceuticals, Inc. (US)				
		Location/Qualifiers				
		1..1863				
		/organism="Homo sapiens"				
		/mol_type="genomic DNA"				
		/db_xref="taxon:9606"				
		1..>1863				
		/note="unnamed protein product"				
		/codon_start=1				
		/protein_id="CAD32050.1"				
		/db_xref="GI:21065828"				
		/translation="MSGGLPLRTAAARACRGLVSTANRLRLTSPVVRFAKELEF				
		LGTKKKEVFPFVPSODELNEINQFLGPVEKFTTEEDSRKIDQEKIGIDETLEKIL				
		SLGLFQVPEEYGGIAGFNTMYRSLGEHISMDGSIITVLAHQAIGLKGILLAGTE				
		QKARYLKLASGEHIAFCLPEPASGSDAASIRKARLSLSEDKKHILNGSKVITNG				
		LANITTFVAKTVEVDSGSKVDKNTLISGRFSMSGVAVGLKKRLIEMTAISACTRKQFNK				
		NTKIPVENILGEVSGDFKVAANILNSGRFSMSGVAVGLKKRLIEMTAISACTRKQFNK				
		RLSFGLEIKFALMAQAYVWESMTYLTAGMDQGPDPDSEIAAMVKVFSSEAA				
		CYSFALQILGLGYTRDYPYERILSDRTLLIFETNEILRMVYIALTGLQHGRLIT				
		RIHELKQAKVSTVMDTVGRRLRDSIGRTVDLGLTGNHGVSHPSRLSADSKNEENTYCE				
		RVFTVLLRLFRGKTIEMERQILKRVANILINLYGMTAVLSRASIIRGLRNHDEHVI				
		LANTFVEAYLQNLFLSLQDKYAPENLDEQIKKVSQQLLEKRAYICAHPLDRTC"				
BASE COUNT		471 a	481 c	519 g	392 t	
ORIGIN						
Query Match		76.0%;	Score 1863;	DB 6;	Length 1863;	
Best Local Similarity		100.0%;	Pred. No. 0;			
Matches 1863;		Conservative	0;	Mismatches	0;	Indels 0; Gaps 0
QY	67	ATGAGCGGTGGGGCTCTCTCTGGCACACCGCTGCGGCTCTGCTCGCGGGGTCTG	126			
Db	1	ATGAGCGGTGGGGGTCTCTCTGGCACACCGCTGCGGCTCTGCTCGCGGGGTCTG	60			
QY	127	GTGGTCTCTACCGGAACCGGGGTCTACTGCGACACGCGCGCTGTAGGAGCTTTCGC	186			
Db	61	GTGGTCTCTACCGGAACCGGGGTCTACTGCGACACGCGCGCTGTAGGAGCTTTCGC	120			
QY	187	AAAGAGCTTTTCTAGGCAAAATCAGAGAAGAAGTCTTCCCATTTCCAGAAGTTAGC	246			
Db	121	AAAGAGCTTTTCTAGGCAAAATCAGAGAAGAAGTCTTCCCATTTCCAGAAGTTAGC	180			
QY	247	CAAGATGAATTAATGAATCAATCAGTTCTTTGGGACCCCGTGGAAAAATTTCTTCACTG	306			
Db	181	CAGATGAATTAATGAATCAATCAGTTCTTTGGGACCCCGTGGAAAAATTTCTTCACTG	240			
QY	307	GAGGTGGACTCCGAAAAATTGACAGGAGGAAAAATCCAGATGAAGCTTTGGAGAAA	366			
Db	241	GAGGTGGACTCCGAAAAATTGACAGGAGGAAAAATCCAGATGAAGCTTTGGAGAAA	300			
QY	367	TTCAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCCAGAGAAATATGGTGGCCTGGGGCTT	426			
Db	301	TTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCCAGAGAAATATGGTGGCCTGGGGCTT	360			
QY	427	TCCAACACCATGTACTCAAGCTAGGGGAGATCATCAGCATGATGGTGCATCACTGTG	486			
Db	361	TCCAACACCATGTACTCAAGCTAGGGGAGATCATCAGCATGATGGTGCATCACTGTG	420			

QY 487 ACCCTGGCAGCGCACCAGGCTATTGGCTCAAGGGATCATCTTGGCTGGCAGTGGAG 546
 Db 421 ACCCTGGCAGCGCACCAGGCTATTGGCTCAAGGGATCATCTTGGCTGGCAGTGGAG 480
 QY 547 CAGAAAGCAAAATCTTGGCTTAAACTGGCTGCGGGAGGACACATTGGCAGCTTCTGGCTC 606
 Db 481 CAGAAAGCAAAATCTTGGCTTAAACTGGCTGCGGGAGGACACATTGGCAGCTTCTGGCTC 540
 QY 607 ACGGAGCAGCAGTGGGAGGATGCGAGCTCAATCCGGAGCAGAGCCACACTAAGTGAA 666
 Db 541 ACGGAGCAGCAGTGGGAGGATGCGAGCTCAATCCGGAGCAGAGCCACACTAAGTGAA 600
 QY 667 GACAAGAGCACTACATCTCAATGGCTCCAAGTCTGGATTACTAATGAGGACTGGCC 726
 Db 601 GACAAGAGCACTACATCTCAATGGCTCCAAGTCTGGATTACTAATGAGGACTGGCC 660
 QY 727 AATATTTTACTGTGTTGCAAGACTGAGGTCTGGATTCTGATGATGATGATGATGATG 786
 Db 661 AATATTTTACTGTGTTGCAAGACTGAGGTCTGGATTCTGATGATGATGATGATGATG 720
 QY 787 AAAATCACAGCACTTATAGTAGAAGAGACTTGGTGGAGTCACTAATGGGAACCCGAA 846
 Db 721 AAAATCACAGCACTTATAGTAGAAGAGACTTGGTGGAGTCACTAATGGGAACCCGAA 780
 QY 847 GATAAATAGGCATTGGGGCTCCAACTGTGGAAGTCCATTTTGAAGAACCAAGATA 906
 Db 781 GATAAATAGGCATTGGGGCTCCAACTGTGGAAGTCCATTTTGAAGAACCAAGATA 840
 QY 907 CCTGTGGAAGCACTCTTGGAGAGTGGGATGGGTTTAAAGTGGCCATGACATCTCTC 966
 Db 841 CCTGTGGAAGCACTCTTGGAGAGTGGGATGGGTTTAAAGTGGCCATGACATCTCTC 900
 QY 967 AACAGCGCGCGTTCAGCATGGCAGCGCTGCTGGCTGGCTGCTCAAGAGATTGATTGAA 1026
 Db 901 AACAGCGCGCGTTCAGCATGGCAGCGCTGCTGGCTGGCTGCTCAAGAGATTGATTGAA 960
 QY 1027 ATGACTGCTCAGTACACAGCAAGCAAGTAAAGAGGCTCAGTGAATTTGA 1086
 Db 961 ATGACTGCTCAGTACACAGCAAGCAAGTAAAGAGGCTCAGTGAATTTGA 1020
 QY 1087 TTGATTTCAGGAGAAATTTGCATGATGGCTCAGAGGCTTACGTATGAGATGATGACC 1146
 Db 1021 TTGATTTCAGGAGAAATTTGCATGATGGCTCAGAGGCTTACGTATGAGATGATGACC 1080
 QY 1147 TACCTCACAGCAGGATCTGACACACCTGCTTCCCGACTGCTCCATCGAGCAGCC 1206
 Db 1081 TACCTCACAGCAGGATCTGACACACCTGCTTCCCGACTGCTCCATCGAGCAGCC 1140
 QY 1207 ATGGTGAAGGTGTTGAGTCCGAGCGCGCTGCGCAGTGTGAGTGAGGCGCTGCAGATC 1266
 Db 1141 ATGGTGAAGGTGTTGAGTCCGAGCGCGCTGCGCAGTGTGAGTGAGGCGCTGCAGATC 1200
 QY 1267 CTCGGGGGTTGGCTTACACAGGAGCTATCGGTACGAGCGCATACTCGGTGACACCGC 1326
 Db 1201 CTCGGGGGTTGGCTTACACAGGAGCTATCGGTACGAGCGCATACTCGGTGACACCGC 1260
 QY 1327 ATCTCTCTCATCTTCAGGAGCAACATGAGATTCTCCGGATGATACATCGCCCTGACGGGT 1386
 Db 1261 ATCTCTCTCATCTTCAGGAGCAACATGAGATTCTCCGGATGATACATCGCCCTGACGGGT 1320
 QY 1387 CTGAGCATGCGGGCGCATCTGACTACCAAGGATCCATGAGCTTAAACAGCCCAAGTG 1446
 Db 1321 CTGAGCATGCGGGCGCATCTGACTACCAAGGATCCATGAGCTTAAACAGCCCAAGTG 1380
 QY 1447 AGCAGATCATGATACCGTGGCGGAGGCTTCGGGACTCCCTGGCGCAACTGTGGAC 1506
 Db 1381 AGCAGATCATGATACCGTGGCGGAGGCTTCGGGACTCCCTGGCGCAACTGTGGAC 1440
 QY 1507 CTGGGGCTGACAGGCAACCATGGATTGTGACCCAGCTCTTGGGAGCAGTGCCCAAG 1566
 Db 1441 CTGGGGCTGACAGGCAACCATGGATTGTGACCCAGCTCTTGGGAGCAGTGCCCAAG 1500
 QY 1567 TTGAGGAGAAACCTACTGCTTCGCGCGGACCGCTGGAGACACTGCTGCTCGCTTTGGC 1626

Db 1501 TTTGAGGAGAAACACCTACTGCTTTCGGCGGACCGTGGAGACACTGCTGCTCGCTTTGGC 1560
 QY 1627 AAGACCATCATGAGGAGCAGCTGGTACTGAAGCGGGTGGCAACATCCTCATCAACCTG 1686
 Db 1561 AAGACCATCATGAGGAGCAGCTGGTACTGAAGCGGGTGGCAACATCCTCATCAACCTG 1620
 QY 1687 TATGGCATGACGCGCGCTGCTGCGGGGCGAGCGCTCCATCCGATTTGGGCTCGGCAAC 1746
 Db 1621 TATGGCATGACGCGCGCTGCTGCGGGGCGAGCGCTCCATCCGATTTGGGCTCGGCAAC 1680
 QY 1747 CAGGACCGAGGCTTCTTGGCCAAACACCTTCTGCGTGAAGCTTACTTGCAGATCTC 1806
 Db 1681 CAGGACCGAGGCTTCTTGGCCAAACACCTTCTGCGTGAAGCTTACTTGCAGATCTC 1740
 QY 1807 TTCAGCTCTCTCAGCTGGACAAGTATCTCCAGAAAACCTTAGATGAGCAGATTAAAGAA 1866
 Db 1741 TTCAGCTCTCTCAGCTGGACAAGTATCTCCAGAAAACCTTAGATGAGCAGATTAAAGAA 1800
 QY 1867 GTGTCCCGAGCAGATCTTGTGAGAGCGAGCTTATATCTGTGCCCGCCCTCTGGACAGACA 1926
 Db 1801 GTGTCCCGAGCAGATCTTGTGAGAGCGAGCTTATATCTGTGCCCGCCCTCTGGACAGACA 1860
 QY 1927 TGC 1929
 Db 1861 TGC 1863

BC001817 1494 bp mRNA linear PRI 12-JUL-2001
 Homo sapiens, Similar to NP0002 protein, clone MGC:3368
 IMAGE:3530094, mRNA, complete cds.
 BC001817
 BC001817.1 GI:12804754
 MGC.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Strausberg, R.
 Direct Submission
 Submitted (29-JAN-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 Contact: amadanesystemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
 Greene, Mark Kettman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://limage.llnl.gov>
 Series: IRAL Plate: 11 Row: k Column: 19.
 Location/Qualifiers
 1..1494
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:3368 IMAGE:3530094"
 /tissue_type="Muscle, rhabdomyosarcoma"
 /clone_lib="NIH_MGC_17"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 475..1041

FEATURES
 source
 1..1494

CDS

/codon_start=1		850		CGAACACGACGACGAGGTTCTTCTGGCCACACACCTTCTGCTGGAGCTTACTTGCAG		909	
/product="Similar to NP0002 protein"		1801		AATCTCTTACGCTCTCTCAGCTGGACAGATATGCTCCAGAAACCTAGATGACAGATT		1860	
/protein_id="AAH01817.1"		910		AATCTCTTACGCTCTCTCAGCTGGACAGATATGCTCCAGAAACCTAGATGACAGATT		969	
/db_xref="GI:12804755"		1861		AAGAAAGTGTCCAGCAGATCTTGAAGAGGAGCCTATATCTGTGCCACCTCTGCGAC		1920	
translation="MYALTGLQHGRIITRIHELKQAKVSTVMDTVGRRLRSLGR		970		AAGAAAGTGTCCAGCAGATCTTGAAGAGGAGCCTATATCTGTGCCACCTCTGCGAC		1029	
TVDLGLTGNHGVHPVLSADSNKFNFTYCFRTVETLILRFKGT IMEQVLKRVAN		1921		AGGACATGCTGAGCAGGAGGACAGTCCCTGCTAGCGCGCGCCCTACCCATGGCCCG		1980	
ILNLGMYAVLSRASRSTRIGLRNHDHEVLANTFCVEAVLQNLFLSLQDKVAPEN		1030		AGGACATGCTGAGCAGGAGGAGCAGTCCCTGCTAGCGCGCGCCCTACCCATGGCCCG		1089	
LDEQIKVQQILEKRAYICAPLDRIC"		1981		TTGCTGGATGACTGTTACTCTTTTTCAGAAAGTGTGGGATTTATACAGGTTAAGCCTT		2040	
BASE COUNT 370 a 386 c 422 g 316 t		1090		TTGCTGGATGACTGTTACTCTTTTTCAGAAAGTGTGGGATTTATACAGGTTAAGCCTT		1149	
ORIGIN		2041		TTGTTCCCGCTGACCTGAAGGTTGTCGCTGGCTGGGAGAGCCTCTTCCAGGTTT		2100	
Query Match 58.7%; Score 1439; DB 9; Length 1494;		1150		TTGTTCCCGCTGACCTGAAGGTTGTCGCTGGCTGGGAGAGCCTCTTCCAGGTTT		1209	
Best Local Similarity 100.0%; Pred. No. 0;		2101		TGACCTGCGAGGAGTCTCTTAACAGGACCATCACAGCTTCTGAACCTGAGCGGAGAGA		2160	
Matches 1439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		1210		TGACCTGCGAGGAGTCTCTTAACAGGACCATCACAGCTTCTGAACCTGAGCGGAGAGA		1269	
901 AAGATACCTGTGAAACATCTTGGAGAGGTCGGAGATGGTTTAAAGTGCCCATGAAC		2161		GAGATGGAATGCTGACCCCTGGAACTGCGGGTATCTGTCTATTCAGGAGACACCAT		2220	
10 AAGATACCTGTGAAACATCTTGGAGAGGTCGGAGATGGTTTAAAGTGCCCATGAAC		1270		GAGATGGAATGCTGACCCCTGGAACTGCGGGTATCTGTCTATTCAGGAGACACCAT		1329	
961 ATCTCAACAGCGCGGTTACAGTGGCGAGCTGCTGGCTGGGCTGCTCAAGAGATTG		2221		AGTGAACCTGGGCTTATGCTGCTCCAGGCTGTGAGTGGGTGGGAGCCTGTGTCTC		2280	
70 ATCTCAACAGCGCGGTTACAGTGGCGAGCTGCTGGCTGGGCTGCTCAAGAGATTG		1330		AGTGAACCTGGGCTTATGCTGCTCCAGGCTGTGAGTGGGTGGGAGCCTGTGTCTC		1389	
1021 ATTGAATGACTGCTGAGTACGCTGCACAAAGAAACAGTTTAAAGAGGCTCAGTGAA		2281		AGGTGGATGACCATTTCTGCTCAACACACATTTCTTAAGAAACAGCTTGAAGCTC		2339	
130 ATTGAATGACTGCTGAGTACGCTGCACAAAGAAACAGTTTAAAGAGGCTCAGTGAA		1390		AGGTGGATGACCATTTCTGCTCAACACACATTTCTTAAGAAACAGCTTGAAGCTC		1448	
1081 TTTGGATGATTACAGGAGAAATTTGCACATGCTGCTCAGAGGCTTACGTCATGGAGAT		RESULT 11					
190 TTTGGATGATTACAGGAGAAATTTGCACATGCTGCTCAGAGGCTTACGTCATGGAGAT		BC033277		2129 bp mRNA linear ROD 16-APR-2003			
1141 ATGACCTACCTCAGCAGGAGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAG		LOCUS		Mus musculus RIKEN cDNA 2600017P15 gene, mRNA (cdna clone MGC:28694			
250 ATGACCTACCTCAGCAGGAGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAG		IMAGE:4241041), complete cds.					
1201 GCAGCATGTTGAAGTGTTCAGTCCAGCGCCCTGGCAGTGTGTGAGTGAAGCGCTG		ACCESSION		BC033277.1 GI:23958794			
310 GCAGCATGTTGAAGTGTTCAGTCCAGCGCCCTGGCAGTGTGTGAGTGAAGCGCTG		VERSION		MGC.			
1261 CAGATCTCCGGGCTTGGGTACACAAAGGACTATCCGTACGAGCGCATACTGCGGTGAC		SOURCE		Mus musculus (house mouse)			
370 CAGATCTCCGGGCTTGGGTACACAAAGGACTATCCGTACGAGCGCATACTGCGGTGAC		KEYWORDS		Mus musculus			
1321 ACCCGCATCTCTCTATCTCGAGGAACCAATGAGATTTCTCCGGATGTACATCGCCCTG		ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
430 ACCCGCATCTCTCTATCTCGAGGAACCAATGAGATTTCTCCGGATGTACATCGCCCTG		REFERENCE		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
1381 ACGGTCTGCAGCATCGCGCGCATCTGACTACAGGATCCATGAGCTTAAACAGGCC		AUTHORS		1 (bases 1 to 2129)			
490 ACGGTCTGCAGCATCGCGCGCATCTGACTACAGGATCCATGAGCTTAAACAGGCC				Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,			
1441 AAGTGAGCACAGTCATGATACCGTTTGGCGGAGGCTTCGGGACTCCCTGGCGGAACT				Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,			
550 AAGTGAGCACAGTCATGATACCGTTTGGCGGAGGCTTCGGGACTCCCTGGCGGAACT				Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,			
1501 GTGGACCTGGGCTGACAGGAACCATGGAGTTGTGACCCAGCTTTCGGAGACAGTGCC				Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,			
610 GTGGACCTGGGCTGACAGGAACCATGGAGTTGTGACCCAGCTTTCGGAGACAGTGCC				Platchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,			
1561 AACAAGTTTGAGGAGAACACCTACTGCTTCCGCGGACCGGTGGAGACACTGCTGCTCGC				Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,			
670 AACAAGTTTGAGGAGAACACCTACTGCTTCCGCGGACCGGTGGAGACACTGCTGCTCGC				Sheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,			
1621 TTTGGACAGACCATCATGAGGAGCAGCTGGTGAAGCGGGTGGCCAACTCTCATC				Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,			
730 TTTGGACAGACCATCATGAGGAGCAGCTGGTGAAGCGGGTGGCCAACTCTCATC				Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,			
1681 AACCTGTATGGCATGAGCGCGCTGCTGTCGGCGGACCGCTCCATCGGATGGGCTC				McKernan, R.D., Malek, J.A., Gunaratne, P.H., Richards, S.,			
790 AACCTGTATGGCATGAGCGCGCTGCTGTCGGCGGACCGGTGGAGACACTGCTGCTCGC				Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,			
1741 CGCAACACGACACAGAGTTCTTCTGGCCAAACACCTTCTGCTGGAGCTTACTTGCAG				Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,			
				Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,			
				Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,			
				Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,			
				Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,			
				Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,			
				Schnerch, A., Schein, J.E., Jones, S.J. and Marra, W.A.,			
				Generation and initial analysis of more than 15,000 full-length			
				human and mouse cDNA sequences			
				Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			

1238 GGACGTGTGTAGTGGAGCGCTGCAGATCCTCGGGGCTTGGCTACCAAGGACTATC 1297
1214 GGCAGTGTGTAGCGAGGCTCTGCAGATCCTTGGGGGCTCAGCGATACATGAGGACTACC 1273
1298 CGTACGAGCGCATACCTGCGTGCACACCGCATCCTCTCATCTTCGAGGGAACCAATGAGA 1357
1274 CCTATGAGCGCATCTGCGTGTGATGCCGATCTCTCTATCTTTGAGGGAACCAATGAGA 1333
1358 TTCTCCGGATGTACATGCGCCGTGACGGGTCTGCAGCATCGCGGCCGACCTCTTGACTACCA 1417
1334 TTCTCCGGTGTTCATTTGGCTGCACAGGCTGCAGCATCTGCAGCATCTCTGACTCGA 1393
1418 GGATCATGAGCTTAACAGGCGCAAGTGAGCACAGTGCATGATACCGTTGGCCGAGGC 1477
1394 GGATCAAGAGCTTAAAGTGGCAATGTGACCAAGTGCATGAGAGCATGTGTCGGAAC 1453
1478 TTGCGGACTCCCTGGGCGCAACTGTGGACCTGGGCTGCAGCAACCAACCATGGATGTGTC 1537
1454 TTGCGGATCTCTGGGCGCAACTGTGGACCTGGGCTTAACAGCGCATCTTGGTGTGTGCC 1513
1538 ACCCGAGTCTTGGGACAGTGCACAAAGTTTGAGGAGACACCTACTGCTCGCGCGCA 1597
1514 ACCCGAGCTTGGCGAGTGCACAAAGCTTGAGGAAATGTCCATTACTTTGGCCGGA 1573
1598 CCGTGGAGACACTGCTCCGCTTTGGCAAGACCATCATGAGGAGCAGCTGGTACTGA 1657
1574 CTGTGTAGACCTGTACTCCGTTTGGAAAGAACATGTGAGAGCAACAGCTGGTGTAA 1633
1658 AGCGGTGGCAACATCCTCATCACTGTATGATGATGAGCGCGCTGCTCGCGGCCA 1717
1634 AGCGGTGAGCAACATCCTCATCACTGTATGATGATGAGCTGTGCTGTCGAGAGCCA 1693
1718 GCCCTCCATCCGATTTGGCTCGCAACACGACGAGGTTCTTTGGCCCAACACCT 1777
1694 GCCCTCCATCAGAAATTTGGCTCAGGATCATGATCAGAGGTTCTTTGGCCCAACATGT 1753
1778 TCTCGTGGAGCTTACTTGCAGAACTCTTTCAGCCCTCTCTCAGCTGGCAAGTATGCTC 1837
1754 TCTGTGTGGAAGCTTATTTCCAGAACTCTTTCAGCCCTGTCTCAGCTGGCAAGAGCTC 1813
1838 CAGAAACCTTATGAGCAGATTAAGAAAGTGTCCAGCAGATCTCTTGAAGAGGAGCT 1897
1814 CAGAAACCTTATGAGCAGATTAAGAAAGTGTCTCGCAGATCTCTTGAAGAGGAGCT 1873
1898 ATATCTGTGCCACCCCTCTGGACAGGACATGCTGAGCAGGAGCA 1943
1874 ATATCTGTGCCACCCCTGGACAGGACATCTCTGAGGCTGGAACCA 1919

RESULT 12
BC032213
LOCUS
DEFINITION Mus musculus RIKEN cDNA 2600017P15 gene, mRNA (cDNA clone MGC:38234 IMAGE:5323723), complete cds.
ACCESSION BC032213
VERSION BC032213.1 GI:21595208
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 2148)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, J., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettner, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnerfeld, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 2148)
Strausberg, R.
Direct Submission
Submitted (06-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-re@mail.nih.gov
Tissue procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

REMARK
COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 56 Row: 1 Column: 12
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis. Location/Qualifiers
1. .2148
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:38234 IMAGE:5323723"
/tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI-CGAP_Mam6"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1. .2148
/gene="2600017P15Rik"
/note="synonym: 4732402K02"
/db_xref="LocusID:229211"
/db_xref="MGI:1914272"
51. .1928
/codon_start=1
/product="RIKEN cDNA 2600017P15"
/protein_id="AAH32213.1"
/db_xref="GI:21595209"
/db_xref="LocusID:229211"
/translation="MSGCVLLSRGATAAAARARVLRFFTRRRPLHTSLQSCSFA KFLPLGNKQKQVFPFVQSHLSSEINQVPLKFFFTFVDSRDKDQKIPVDTL EKLSGLFGIOVPEYGGIGLSTMYRLGEISLDSITVTIAHQALGKILIV GNDQKAKYLPKLSSGHEHIAAFCLETPASGSDAASIQTRATLSEDKKYFLNGSKYMI

FEATURES
source

gene

CDS

TNGLIANIFTVAKTEVVVDSGSKDKMTAFIVERDFGINTGPKDKLIGRNTCE
VHENTRVPVENLVEGVGKFKVAMNINLSGRFSNGSAVAGMLKLLIELTAETARX
QPNRLNLSFGLIOEKFAIMAKOAVYMESMAYLTSOMLDQPGFPCDSIEAAVMKVFSE
AAWOCVSEALQILAGSGYMKDYPFERMLRDKARILLIFEETNEILRFLATGQHQAGR
LLTSRIKELKSGNVTVMETIGRKLRSLSLGTDLGTVGVVHSLGSDSANKLEH
VHYFERTVETLLLRFGKNIVEQJLVLRVANILINLYGMTAVLSRASRSIRIGLRNH
HEVLIANMFCVEAFYQNLFSLSQLDKNAPENLDQIKKVSRIQLEKRAYICAPLDR
S*

BASE COUNT 554 a 504 c 593 g 497 t
ORIGIN

Query Match 56.6%; Score 1388.4; DB 10; Length 2148;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 1560; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY	98	CGGCTGCGGCTCGTCCCTGCGCGGCTCTGTTGGTCTTACCGCGAAACCGCGGCTACTGC	157
DB	94	CGGCTGCGGCGCGAGCTCCCGGCTGCTGAGGGAATTCACGGCGCGCGCGCGCTGC	153
QY	158	GCACCGCCGCTGTAGGAGCTTTCGCCAAAGAGCTTTCTAGGCAAAATCAAGAAGA	217
DB	154	ACACAGCCTGACAGAGCTGCTGTTTCGCCAAGGAGCTTCTCCTGGGCAACATCAAGCAGA	213
QY	218	AAGAAGTTTCCCAATTTCCAGAGCTTAGCCAAAGATGAACCTATATGAATCAATGATCT	277
DB	214	AAGAGTTTCCCAATTTCCAGAGCTAGTCAACATGAACCTTAGTGAATCAATCACTGCTG	273
QY	278	TGGACCGCTGGAAAAATTTCTACTGAAGAGGTGGACTCCCGAAAAATTTGACCAAGAAG	337
DB	274	TGGACCGCTGGAAAAATTTCTACTGAAGAGGTGGACTCCCGAAAAATTTGACCAAGAAG	333
QY	338	GAAAAATCCAGATGAACCTTTGGAGAAATTTGAAGAGCTTAGGGCTTTTGGGCTGCAAG	397
DB	334	GAANAATCCAGTTTGACACCTTAGAAGTTAGAGAGCTGGGACTTTTGGCATACAGG	393
QY	398	TCCAGAAATATGTGGCTGGGCTTCTCCTCAACACCATGTACTCAAGACTAGGGGAGA	457
DB	394	TCCAGAAATATGTGGCTGGGCTTCTCCTCAACACCATGTACTCAAGACTAGGGGAGA	453
QY	458	TCATCAGATGATGGGTGCTACATGACCTGACCGTGGCAGCAGCAGGCTATCGGCCCTCA	517
DB	454	TCATCAGCTGGATGGCTCTATACGCTGACCTGGCAGCAGCAGGCTATCGGCCCTCA	513
QY	518	AGGGATCATCTTGGCTGGGCTGAGGAGCAGAAAGCCAAATACTTGCTTAACCTGGGT	577
DB	514	AGGGATCATCTTGGCTGGGCTGAGGAGCAGAAAGCCAAATACTTGCTTAACCTGGGT	573
QY	578	CCGGGAGCACATTTGAGGCTTCTGCTCACGGAGCCAGCCAGTGGGAGCATGACGCT	637
DB	574	CTGGGAACACATCGCAGCTTCTGCTAACAGAGCCAGCAGTGGGAGTATGCTGCAT	633
QY	638	CAATCCGGAGCAGCCACACTAAGTGAAGACAGAAGACACTACATCTCAATGGCTCCA	697
DB	634	CAATCCAGACAGAGCTAGTGAAGTGAAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	693
QY	698	AGTCTGGATTACTAATGAGGACTCGGCAATATTTTACTGTGTTTCAAGACTGAGG	757
DB	694	AGTCTGGATTACTAATGAGGACTCGGCAATATTTTACTGTGTTTCAAGACTGAGG	753
QY	758	TGCTTGATCTCATGGATGAGTGAAGACAAAATCACAGCATTCATAGTAGAAGAGACT	817
DB	754	TGCTTGATCTCATGGATGAGTGAAGACAAAATCACAGCATTCATAGTAGAAGAGACT	813
QY	818	TTGTTGGAGTCACTAATGGAAACCCGAAGATTAATTAGGATTTCCGGGCTCCCAACTT	877
DB	814	TCGGCGGAATCACTAATGGAAACCCGAAGATTAATTAGGATTTCCGGGCTCCCAACTT	873
QY	878	GTGAAGTCCATTTTGAATAACCAAGATACCTGTGAAAAATCTCTTGGAGAGCTCGAG	937
DB	874	GTGAAGTCCATTTTGAATAACCAAGATACCTGTGAAAAATCTCTTGGAGAGCTCGAG	933
QY	938	ATGGGTTTAAGTGGCCATGAACATCTTCAACAGCGCGGCTTACAGCATGGCAGCGCTG	997
DB		ATGGGTTTAAGTGGCCATGAACATCTTCAACAGCGCGGCTTACAGCATGGCAGCGCTG	

DB	934	GGGCTTTTAGGTAGCCATGAACATCTTGAACAGTGGCGAATTCAGCATGGGCGTGTG	993
QY	998	TGCTGGCTGGCTCAAGAGATGATTAATGAACTGCTGAGTACGCTTCAAGCAAGAAC	1057
DB	994	TGCGCGGGATGCTGAAGAAGCTGATGAACATGACTGCTGAGTATGCTTACGAGAAC	1053
QY	1058	AGTTTAAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGGACATGATGGCTC	1117
DB	1054	AGTTTAAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGGACATGATGGCTC	1113
QY	1118	AGAAGCTTACCTCATGGAGAGATATGACCTTACCTCAGACAGGAGATGCTGGACCACTG	1177
DB	1114	AGAAGCTTACCTCATGGAGAGATATGACCTTACCTCAGACAGGAGATGCTGGACCACTG	1173
QY	1178	GCTTTCCGACTGCTCCATTCAGGAGGAGCATGCTCAAGGTGTTCAGCTCCGAGCGGCT	1237
DB	1174	GATTTCCGACTGCTCTATTCAGGAGGAGCATGCTCAAGGTGTTCAGCTCCGAGCGGCT	1233
QY	1238	GGCAGTGTGTGAGTGAAGGCTGAGATCTCGGGGGCTTGGGCTACACAGGAGCTATC	1297
DB	1234	GGCAGTGTGTGAGGAGGCTCTGCAAGTCTTGGGGGCTCAGGCTTACATGAAGACTACC	1293
QY	1298	CGTACGAGCGCATACTGCTGACACCCGCACTCTCTCATCTTTCAGGAGAACCAATGAGA	1357
DB	1294	CCTATGAGCGCATGCTGCTGATGCCGCACTTCTCTAATCTTTGAGGAGAACCAATGAGA	1353
QY	1358	TTCTCCGATGTACATCCCTGAGCGGTCTGCAGCATCCGCGGCTGACATCTGACTACCA	1417
DB	1354	TTCTTCGGTGTTCATTCGCTGACAGGCTTGCAGCATCTTGGGGCTGACATCTGACTACCA	1413
QY	1418	GGATCCATGAGCTTAAACAGGCGCAAGTGAACACAGATCTATGATACCTGTCGCGAGGC	1477
DB	1414	GGATCAAGAGCTTAAAGTGGCAATGTGACACAGTCTGAGGAGAACCAATGAGA	1473
QY	1478	TTGCGGACTCCCTGGGCGGAACTGTGACCTTGGGCTGACAGGCAACCATGAGTGTCTGC	1537
DB	1474	TTGCGGATCTCTGGGCGGAACTGTGACCTTGGGCTTACAGGCGCATCTTGTGTGTGTC	1533
QY	1538	ACCCAGCTTTCGCGGAGTCCCAAGATTTGAGGAGAACACCTACTGCTTCGCGCGGA	1597
DB	1534	ACCCAGCTTTCGCGGAGTCCCAAGATTTGAGGAGAACACCTACTGCTTCGCGCGGA	1593
QY	1598	CCGTGGAGACACTGCTGCTTCCGCTTGGCAAGACCATCATGGAGGAGCACTGCTACTGA	1657
DB	1594	CTGTTGAGACCTGTTACTCCGCTTGGAAAGACATTTGAGGAGAACCACTGGGTAA	1653
QY	1658	AGCGGTGGCGAACATCTCATCAACCTGTATGGCATGACGCGGCTGCTGCTCGCGGCA	1717
DB	1654	AGCGGTGGCGAACATCTCATCAACCTGTATGGCATGACGCGGCTGCTGCTCGCGGCA	1713
QY	1718	GCGCTCCATCCGCTTGGGCTCCGCAACCCAGCAGGATCTCTTGGGCAACACCT	1777
DB	1714	GCGCTCCATCCGCTTGGGCTCCGCAACCCAGCAGGATCTCTTGGGCAACACCT	1773
QY	1778	TCTGCTGGAAGCTTACTTGCAGATCTCTTACGCTCTCTCAGCTGGACAGTATGCTC	1837
DB	1774	TCTGCTGGAAGCTTACTTGCAGATCTCTTACGCTCTCTCAGCTGGACAGTATGCTC	1833
QY	1838	CAGAAACCTTAGATGAGCAGATTAAGAAAGTGTCCAGCAGATCTCTTGAAGACGAGCT	1897
DB	1834	CAGAAACCTTAGATGAGCAGATTAAGAAAGTGTCTCGCAGATCTCTTGAAGACGAGCT	1893
QY	1898	ATATCTGTGCCACCTCTGGCAGCAGCATGCTGAGGCGGGGACA	1943
DB	1894	ATATCTGTGCCACCTCTGGCAGCAGCATGCTGAGGCGGGGACA	1939

RESULT 13
BC031137
LOCUS
DEFINITION
IMAGE:4971848), complete cds.
ACCESSION
BC031137

3071 bp mRNA linear ROD 10-JUN-2003
Mus musculus RIKEN cDNA 260007P15 gene, mRNA (cdna clone MGC:37218
IMAGE:4971848), complete cds.
BC031137

Db	809	TCGGCGGAATCATAATATGGAAACCCGAGATAAAATTAGGCATTTCGAGGCTCCAACACAT	868
Qy	878	GTGAAGTCCATTTTGA AAAACACAAAGATACCTGTGGA AAAACATCCTTTGGAGAGCTCGGAG	937
Db	869	GTGAGGTCCATTTTGA AATACCAAGATGCCCTGTGGAAATGTCTTGGAGAAGTTGGAG	928
Qy	938	ATGGGTTTAAAGTGGCCATGAACATCCTCAACAGCGCGGTTTCAGCATGGCAGCGCTCG	997
Db	929	GGGCTTTTAAAGTAGCCATGAACATCCTCAACAGTGGCGGATTTCAGCATGGCAGTGC	988
Qy	998	TGCTGGGCTGTCTCAAGAGATTGATTGAAATGACTGCTGAGTACGCCCTGCACAGGAAC	1057
Db	989	TGSCCGGATGCTGAAGAGCTGATCGAACTGACTGCTGAGTATGCTCTTACGAGGAAC	1048
Qy	1058	AGTTTAAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCACATGATGGCTC	1117
Db	1049	AGTTTAAAGAGGAACTCAGTGAATTTGGGTTGATTCAGGAGAAATTTGGCTCATGGCTC	1108
Qy	1118	AGAAGGCTTACGTGCATGAGAGATTGACTACCTACCTACAGCAGGAGATGCTGGACCAACCTG	117
Db	1109	AGAAGCATACGTAAATGAGAGATTGGCTTACCTCACCTCAGGAGATGCTAGACCAACCAAG	1168
Qy	1178	GCCTTCCGACTGCTCCATCGAGCAGCCATGCTGGAAGGTGTTTCAGCTCCGAGGCGCGCT	1237
Db	1169	GATTTCCGACTGCTCTATCGAGCAGCCATGTTGGAAGTGTTCAGCTCTGAGGCTGCCCT	1228
Qy	1238	GGCAGTGTGAGTGAGGCGCTGCAGATCCTCGGGGCTTGGGCTACAAAGGACTATC	1297
Db	1229	GGCAGTGTGAGGAGGCTCTGCAGATCCTTGGGGCTCAGGCTACATGAAGAGCTAC	1288
Qy	1298	CGTAGCAGCGCATACTCGGTGACACCCGCGCATCTCTCTCATCTTCAGAGGAACCAATGAGA	1357
Db	1289	CGTATGAGCGCATGCTGGTGATGCCCGCATCTCTCTATCTTTGAGGGAACAATGAGA	1348
Qy	1358	TTCTCCGATGATGATCGCCCTGACGGGCTGCAGCATGCCGCGCATCCTGACTACCA	1417
Db	1349	TTCTTCGGTGTTCATTTCGGCTGACAGGCGCTGCAGCATGCTGGAGCATCCTGACTCGA	1408
Qy	1418	GGATCCATGAGCTTTAAACAGCGCCAAAGTAGACACAGTCATGATACCGTTGCCGCGAGC	1477
Db	1409	GGATCAAGAGCTTTAAAGTGGCAATGTGACCACATCATGAGACGATTGTGCGGAAC	1468
Qy	1478	TTCCGGACTCCTTGGGCGCACTGTGACCTGGGGCTGACAGCAACCATGGAGTTGTGC	1537
Db	1469	TTCCGGATTCTTGGGCGCAACTGTGGAGCTGGGGCTAACAGCGCATCTTGTGTGTGCTC	1528
Qy	1538	ACCCAGCTTGTGGGACAGTGCCCAACAAGTTTGAGGAGAACCTACTGCTTCGCCCGGA	1597
Db	1529	ACCCAGCTTGGGACAGTGCCCAACAAGCTTGAGGAATGTCCATTACTTTTGGCGCGGA	1588
Qy	1598	CCGTGAGACACTGCTGCTCCGTTTGGCAAGACCATCATGAGGAGGACAGCTGTACTGA	1657
Db	1589	CTGTTGAGACCTGTACTCCGTTTGGAAAGAAATTTGAGGAACAGCTGGTGTAA	1648
Qy	1658	AGCGGTGGCCAACTCCTCATCAACCTGATGGGATGAGCGGCTGTCTTGGCCCAACCT	1717
Db	1649	AGCGGTGAGCAACATCTCATCACTGATGGATGAGCGGCTGTCTGTCGCGGGCCA	1708
Qy	1718	GGCGCTCCATCCGATTTGGGCTCCGAACTCACTGATGGATGACTGCTGTGTGTCGAGAGCCA	1777
Db	1709	GGCGCTCCATCAGAAATTTGGGCTCAGGAATCATGATCAGCAGGTTCTGTGTGCCAACATGT	1768
Qy	1778	TCGCTGGAAGCTTACTTGCAGAACTCTTCAGGCTCTCTCAGCTGCACAAGTATGCTC	1837
Db	1769	TCGTGTGGAAGCTTATTTCCAGAACTCTTCAGGCTGTCTCAGCTGGCAAGAACGCTC	1828
Qy	1838	CAGAAACCTTAGATGAGCAGATTAAAGAAAGTGTCCAGCAGATCCTTTGAGAAGCGGCT	1897
Db	1829	CAGAAACCTTGGATGAGCAGATTAAAGAAAGTGTCTCGCAGATCCTTTGAGAGCGGCT	1888
Qy	1898	ATATCTGCCCCACCTCTGGACAGGACATGCTGAGGAGGGGACA	1943
Db	1889	ATATCTGCCCCACCTCTGGACAGGACATCCTGAGGCTGGAAGCA	1934

RESULT 14	AK027186	1995 bp	linear	PRI 29-SEP-2000
LOCUS	AK027186			
DEFINITION	Homo sapiens CDNA: FLJ23533 fis, clone LNC06931, highly similar to AF078654 Homo sapiens NP0002 mRNA.			
ACCESSION	AK027186			
VERSION	AK027186.1 GI:10440254			
KEYWORDS	olligo capping; fis (full insert sequence).			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 (sites)			
	Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M., Ohmori,Y., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.			
TITLE	NEDO human cDNA sequencing project			
JOURNAL	2 (bases 1 to 1995)			
REFERENCE	Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.			
AUTHORS	Direct Submission			
TITLE	Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp. Tel:81-3-5449-5286, Fax:81-3-5449-5416)			
JOURNAL	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert construction, 5'- & 3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).			
COMMENT	Location/Qualifiers 1..1995 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="LNC06931" /tissue_type="human lung" /clone_lib="LNC" /note="cloning vector pME18SFL3" 1..1995 /note="highly similar to AF078654 Homo sapiens NP0002 mRNA"			
FEATURES	misc_feature			
BASE COUNT	492 a 513 c 553 g 437 t			
ORIGIN				
Query Match	31.0%; Score 759.6; DB 9; Length 1995;			
Best Local Similarity	99.0%; Pred. No. 1.7e-177;			
Matches 796;	Conservative 0; Mismatches 4; Indels 4; Gaps 3;			
QY	1626 CAAGACCATCATGAGGAGCAGCTGGTACTGAAGCGGGTGGCCAAACCTTCATCAACCT 1685			
Db	1183 CCAGACCATCATGAGGAGCAGCTGGTACTGAAGCGGGTGGCCAAACCTTCATCAACCT 1242			
QY	1686 GTATGGCATGACGGCGGTGCTGTCGGGGGCGAGCC-GCTCCATCCGCTGGGCTCCGCA 1744			
Db	1243 GTATGGCATGACGGCGGTGCTGTCGGGGGCGAGCCAGCTCCATCCGCTGGGCTCCGCA 1302			
QY	1745 ACCAGACCAGAGGTTCTCTTGGCCAAACCTTCTTCGCTGGAAGCTTACTTGCAGAATC 1804			
Db	1303 ACCAGACCAGAGGTTCTCTTGGCCAAACCTTCTTCGCTGGAAGCTTACTTGCAGAATC 1362			
QY	1805 TCTTCAGCCTCTCTCAGCTCGACAAGTATGCTCCAGAAAACCTTAGATGAGCAGATTAA 1864			
Db	1363 TCTTCAGCCTCTCTCAGCTCGACAAGTATGCTCCAGAAAACCTTAGATGAGCAGATTAA 1422			

FEATURES		Location/Qualifiers		/organism='Homo sapiens (human)'	
source		1..771			
BASE COUNT		204 a 187 c 219 g 158 t		3 others	
ORIGIN					
Query Match		30.6%; Score 750.6; DB 6; Length 771;			
Best Local Similarity		99.08; Pred. No. 2.6e-175;			
Matches 764; Conservative 0; Mismatches 7; Indels 1; Gaps 1;					
QY	21	GCTAAGAGGGGAGACTGAGCGCTGAGCGCTGGGGAACATCGGCAGCATGAGCGGTCGGG	80		
DB	1	GCTAAGAGGGGAGACTGAGCGCTGAGCGCTGGGGAACATCGGCAGCATGAGCGGTCGGG	60		
QY	81	GCTCTTCTCGGCACCGCTGCGCTGCTGCGTCCGGGGTCTGGTGGTCTCTACCGC	140		
DB	61	GCTCTTCTCGGCACCGCTGCGCTGCTGCGTCCGGGGTCTGGTGGTCTCTACCGC	120		
QY	141	GAACCGGGGCTACTGCGCACCGCGCTGTACGAGCTTTTCGCAAGAGCTTTTCTCT	200		
DB	121	GAACCGGGGCTACTGCGCACCGCGCTGTACGAGCTTTTCGCAAGAGCTTTTCTCT	180		
QY	201	AGGCAAAATCAAGAAAGAAAGATTTTCCCAATTTCCAGAACTTAGCCAAAGACTTAA	260		
DB	181	AGGCAAAATCAAGAAAGAAAGATTTTCCCAATTTCCAGAACTTAGCCAAAGACTTAA	240		
QY	261	TGAATCAATCAGTTCTTTGGGACCGGTGGAATAATTTCTCACTGAAGAGGTGGACTCCG	320		
DB	241	TGAATCAATCAGTTCTTTGGGACCGGTGGAATAATTTCTCACTGAAGAGGTGGACTCCG	300		
QY	321	AAAAATTGACAGGAGGAAAAATCCAGATGAAACTTTTGAGAAATTTGAAGAGGCTTAGG	380		
DB	301	AAAAATTGACAGGAGGAAAAATCCAGATGAAACTTTTGAGAAATTTGAAGAGGCTTAGG	360		
QY	381	GCTTTTGGGCTGCAAGTCCAGAAAGATATGTTGGGCTGGGCTTCTCCAAACCATGTA	440		
DB	361	GCTTTTGGGCTGCAAGTCCAGAAAGATATGTTGGGCTGGGCTTCTCCAAACCATGTA	420		
QY	441	CTCAAGACTAGGGGAGATCATCAGCATGGATGGTCCATCTACTGTGACCTGGCAGCGCA	500		
DB	421	CTCAAGACTAGGGGAGATCATCAGCATGGATGGTCCATCTACTGTGACCTGGCAGCGCA	480		
QY	501	CCAGGCTATTGGCCTCAAGGGGATCATCTTGGCTGGCACTGAGAGCAGAAAGCCAAATA	560		
DB	481	CCAGGCTATTGGCCTCAAGGGGATCATCTTGGCTGGCACTGAGAGCAGAAAGCCAAATA	540		
QY	561	CTTGCCCTAACTAGGGCTCGGGGGAGCACATTTCAGGCTTCTGCTCAGGAGCCAGCCAG	620		
DB	541	CTTGCCCTAACTAGGGCTCGGGGGAGCACATTTCAGGCTTCTGCTCAGGAGCCAGCCAG	600		
QY	621	TGGAGCGGATGCGAGCTCAATCCGAGCAGAGCCACACTAAGTGAAGCAAGAAGCACTA	680		
DB	601	TGGAGCGGATGCGAGCTCAATCCGAGCAGAGCCACACTAAGTGAAGCAAGAAGCACTA	660		
QY	681	CATCTCTCAATGGCTCCCAAGGCTCTGGATTAATGGAGGACTGGCCCAATTTTACTGT	740		
DB	661	CATCTCTCAATGGCTCCCAAGGCTCTGGATTAATGGAGGACTGGCCCAATTTTACTGT	720		
QY	741	GTTTGCNAAGACTGAGGTGTTGATTTCTGATGATGATGATGATGATGATGATGATGAT	792		
DB	721	GTTTGCNAAGACTGAGGTGTTGATTTCTGATGATGATGATGATGATGATGATGATGAT	771		

Search completed: September 6, 2003, 17:23:04
Job time : 9029.27 secs

QY	1865	AAAGTCTCCAGCAGATCCTTGAGAGCGAGCGCTATATCTGTGCCACCCCTCTGCACAGGA	1924		
DB	1423	AAAGTCTCCAGCAGATCCTTGAGAGCGAGCGCTATATCTGTGCCACCCCTCTGCACAGGA	1482		
QY	1925	CATGCTGAGGAGGAGGAGTGTCCCTGCTTACC--GCCCGCCCTACCCATGGCC--GT	1981		
DB	1483	CATGCTGAGGAGGAGGAGTGTCCCTGCTTACC--GCCCGCCCTACCCATGGCC--GT	1542		
QY	1982	TGCTGGATGACGTGTTACTCTTTTTCAGAAAGTGTGGGATATACAGGTTAAGCCCTT	2041		
DB	1543	TGCTGGATGACGTGTTACTCTTTTTCAGAAAGTGTGGGATATACAGGTTAAGCCCTT	1602		
QY	2042	TGTTCCCGCTGTCACCTGAAGGTTGTCGCTGCGCTGGGAGAGCCCTTCCAGGTTT	2101		
DB	1603	TGTTCCCGCTGTCACCTGAAGGTTGTCGCTGCGCTGGGAGAGCCCTTCCAGGTTT	1662		
QY	2102	GACCTCGAGGAGTGTCTTCAACAGGACCATCACAGCTTCTGAAGTGGCGGAGAGAG	2161		
DB	1663	GACCTCGAGGAGTGTCTTCAACAGGACCATCACAGCTTCTGAAGTGGCGGAGAGAG	1722		
QY	2162	AGAAATGGAATGTCGACCCCTGGAAGTGGGCTGATCTGATGAGGAGACACCATTA	2221		
DB	1723	AGAAATGGAATGTCGACCCCTGGAAGTGGGCTGATCTGATGAGGAGACACCATTA	1782		
QY	2222	GTGGAAGTGGGCTGATCTGCTGCTCCAGGCTGAGGTTGGGAGTGGGAGCTGTGTA	2281		
DB	1783	GTGGAAGTGGGCTGATCTGCTGCTCCAGGCTGAGGTTGGGAGTGGGAGCTGTGTA	1842		
QY	2282	GGTGTGGATAGCCATTTCTGCTCAACGACACATCTCTAAGAAAGCTTGAAGCTCG	2341		
DB	1843	GGTGTGGATAGCCATTTCTGCTCAACGACACATCTCTAAGAAAGCTTGAAGCTCG	1902		
QY	2342	TCTGGTCAATCAATTAAGTACAGGAGGACATTAAGCAATGTACCAGGAACCATTT	2401		
DB	1903	TCTGGTCAATCAATTAAGTACAGGAGGACATTAAGCAATGTACCAGGAACCATTT	1962		
QY	2402	AACAAAGATATAAATGTGCACAA	2425		
DB	1963	AACAAAGATATAAATGTGCACAA	1986		
RESULT 15					
BD150623					
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
COMMENT					
1 (bases 1 to 771)					
Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,					
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.					
Primer for synthesizing full-length cDNA and use thereof					
Patent: JP 2002191363-A 5466 09-JUL-2002;					
HELEX RESEARCH INSTITUTE					
OS Homo sapiens (human)					
PN JP 2002191363-A/5466					
PD 09-JUL-2002					
PF 28-JUL-2000 JP 2000280990					
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU					
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,					
PI SAITO,					
PI KEIICHI NAGAI, TETSUJI OTSUKI					
PC					
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10,					
PC C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC					
Primer for synthesizing full-length cDNA and use thereof FH key					
Location/Qualifiers					
1..771					
FT source					

FT /note= "This region is specifically claimed in
FT claim 1 as SEQ ID NO:3"
FT 1933..2452
FT /*tag= C
XX
PN WO200218582-A2.
XX 07-MAR-2002.
XX 31-AUG-2001; 2001WO-US27186.
XX 31-AUG-2000; 2000US-229831P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Meyers R, Hunter JJ;
XX WPI; 2002-329775/36.
XX P-PSDB; AAE21680.
XX
XX New human dehydrogenase polypeptide for diagnosing and treating
XX dehydrogenase-7 associated diseases and disorders e.g. Alzheimer's
XX disease, and to identify modulators of therapeutic use
XX
XX Claim 1; Fig 1; 124pp; English.
XX
XX The invention relates to human dehydrogenase (DHDR)-7 polypeptides
XX referred to as 62112 and nucleic acid molecules encoding such
XX polypeptides. DHDR-7 is an acyl-CoA dehydrogenase, a mitochondrial
XX flavoprotein which catalyses the first step of fatty acid beta-
XX oxidation. Sequences of the invention and their antibodies are
XX useful for treating a disorder, disease or condition which is caused
XX by misregulation (e.g. downregulation or upregulation) of DHDR activity.
XX Examples of disorders include central nervous system (CNS) disorders
XX e.g. Alzheimer's disease (AS), Pick's disease, Parkinson's, Huntington's
XX disease, Gilles de la Tourette's syndrome, multiple sclerosis, epilepsy,
XX amyotrophic lateral sclerosis and Creutzfeldt-Jakob disease; autonomic
XX function disorders e.g. hypertension, depression, schizophrenia, panic
XX migraine, anxiety, obsessive-compulsive disorder and obesity; cardiac
XX related disorder e.g. arteriosclerosis, ischaemia reperfusion injury,
XX restenosis, arterial inflammation, angina, hypertension, cardiomyopathy
XX and arrhythmia; disorders of musculus skeletal system paralysis, muscle
XX weakness e.g. ataxia; cell proliferation, growth, differentiation or
XX migration disorders e.g. cancer, diabetes mellitus, hypothyroidism or
XX hyperthyroidism; reproductive or fertility disorders; autoimmune or
XX or immune deficiency disorders; hepatic disease or dysfunction and
XX metabolic disorders. They are used for screening assays, predictive
XX medicine e.g. diagnostic assays, prognostic assays, monitoring clinical
XX trials, and pharmacogenetics. Polypeptides of the invention are used
XX to identify modulators that modulate their activity. Polynucleotides
XX of the invention are used in gene therapy. The present sequence is
XX cDNA encoding human DHDR-7.
XX
XX Sequence 2452 BP; 616 A; 616 C; 685 G; 535 T; 0 other;
XX
XX Query Match 100.08; Score 2452; DB 24; Length 2452;
XX Best Local Similarity 100.08; Pred. No. 0;
XX Matches 2452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CGTGTGTGTCCTCGCGCGCTAAGAGGGGAGACTGAGCGTGAGGCGGGAACATCG 60
XX
XX 1 CGTGTGTGTCCTCGCGCGCTAAGAGGGGAGACTGAGCGTGAGGCGGGAACATCG 60
XX
XX 61 GGCAGCATGAGCGGTGCGGGCTCTTCCTGCGCACCGACCGGTGCGGCTGCGCGG 120
XX
XX 61 GGCAGCATGAGCGGTGCGGGCTCTTCCTGCGCACCGACCGGTGCGGCTGCGCGG 120
XX
XX 121 GGTCTGTGTCCTTACCGCAACCGCGGCTACTGCGCACCGACCGCTGTACGAGCT 180
XX
XX 121 GGTCTGTGTCCTTACCGCAACCGCGGCTACTGCGCACCGACCGCTGTACGAGCT 180
XX
XX 181 TTCGCCAAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGTTCCTCCATTTCCAGAA 240
XX
XX
XX

181 TTCGCCAAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGTTCCTCCATTTCCAGAA 240
241 GTTAGCCCAAGCTTAATGAAATCAATCACTTCTTGGGACCCGTGGAAAAATTTCTTC 300
241 GTTAGCCCAAGCTTAATGAAATCAATCACTTCTTGGGACCCGTGGAAAAATTTCTTC 300
301 ACTGAAGAGGTGGATCCCGAAAAATGACCAAGAGGGAATCCAGATGAACCTTTC 360
301 ACTGAAGAGGTGGATCCCGAAAAATGACCAAGAGGGAATCCAGATGAACCTTTC 360
361 GAGAAATGAGAGCCTTAGGCTTTTGGGCTGCAAGTCCCAAGAAATATGCTGGCCTG 420
361 GAGAAATGAGAGCCTTAGGCTTTTGGGCTGCAAGTCCCAAGAAATATGCTGGCCTG 420
421 GGCCTTCCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGTCCATC 480
421 GGCCTTCCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGTCCATC 480
481 ACTGTGACCTTGGCAGCGCAGGCTATTGGCCCTCAAGGGGATCATCTTTGGCTGGCAGCT 540
481 ACTGTGACCTTGGCAGCGCAGGCTATTGGCCCTCAAGGGGATCATCTTTGGCTGGCAGCT 540
541 GAGGAGCAAAAGCAAAATACTTGCCTAAACTGGCGTCCGGGGAGACATTTGCAAGCCTTC 600
541 GAGGAGCAAAAGCAAAATACTTGCCTAAACTGGCGTCCGGGGAGACATTTGCAAGCCTTC 600
601 TGCCTTACGAGCCAGCGCAGTGGGAGCGATCGACCTCAATCCGAGCAGAGCCACACTA 660
601 TGCCTTACGAGCCAGCGCAGTGGGAGCGATCGACCTCAATCCGAGCAGAGCCACACTA 660
661 AGTGAAGCAAAAGCAAAATACTTGCCTAAACTGGCGTCCGGGGAGACATTTGCAAGCCTTC 720
661 AGTGAAGCAAAAGCAAAATACTTGCCTAAACTGGCGTCCGGGGAGACATTTGCAAGCCTTC 720
721 CTGGCCCAATATTTTACTGTGTTCGCAAGACTGAGGTCGTTGATCTCTGATGGATCAGTG 780
721 CTGGCCCAATATTTTACTGTGTTCGCAAGACTGAGGTCGTTGATCTCTGATGGATCAGTG 780
781 AAAGCAAAATCAAGCAAAATCTATAGAAAGAGACTTTGGTGGAGTCACTAATGGGAAA 840
781 AAAGCAAAATCAAGCAAAATCTATAGAAAGAGACTTTGGTGGAGTCACTAATGGGAAA 840
841 CCCGAAGATAAATAGGCAATTCGGGCTCCAAACACTTGTGAAGTCCATTTTGAAGAACACC 900
841 CCCGAAGATAAATAGGCAATTCGGGCTCCAAACACTTGTGAAGTCCATTTTGAAGAACACC 900
901 AAGATACCTGTGGAAGCAATCTTGGAGAGTTCGGAGATGGGTTTAAAGTGCCCATGAAC 960
901 AAGATACCTGTGGAAGCAATCTTGGAGAGTTCGGAGATGGGTTTAAAGTGCCCATGAAC 960
961 ATCCTCAACAGCGCGGCTTTCAGCATGGGCGCTGCGGCTGGGCTGCTCAAGAGATTG 1020
961 ATCCTCAACAGCGCGGCTTTCAGCATGGGCGCTGCGGCTGGGCTGCTCAAGAGATTG 1020
1021 ATTGAATGACTGCTGAGTACGCTGCACAAGGAAACACTTTTAAAGAGGCTCAGTGAA 1080
1021 ATTGAATGACTGCTGAGTACGCTGCACAAGGAAACACTTTTAAAGAGGCTCAGTGAA 1080
1081 TTTGGATTGATTCAGAGAAATTTGCACTGTGGCTCAGAAGGCTTACGTCATGAGAGT 1140
1081 TTTGGATTGATTCAGAGAAATTTGCACTGTGGCTCAGAAGGCTTACGTCATGAGAGT 1140
1141 ATGACCTTACCTCAGCAGGAGTGTGGACCAACCTTGGCTTCCGACCTGCTCCATCGAG 1200
1141 ATGACCTTACCTCAGCAGGAGTGTGGACCAACCTTGGCTTCCGACCTGCTCCATCGAG 1200
1201 GCAGCCATGCTGAAGTGTTCAGCTCCGAGCGCGCTGGCAGTGTGTGAGTGAGGCGCTG 1260
1201 GCAGCCATGCTGAAGTGTTCAGCTCCGAGCGCGCTGGCAGTGTGTGAGTGAGGCGCTG 1260
1261 CAGATCTCGGGGCTTGGCTACAAAGGAGATTCCTGACAGCGCATCTGCTGCTGAC 1320
1261 CAGATCTCGGGGCTTGGCTACAAAGGAGATTCCTGACAGCGCATCTGCTGCTGAC 1320


```
QY 1321 ACCCGCATCTCTCATCTTCGAGGAACCAATGAGATTCTCCGGATGTACATCGCCCTG 1380
Db 1321 ACCCGCATCTCTCATCTTCGAGGAACCAATGAGATTCTCCGGATGTACATCGCCCTG 1380
QY 1381 ACGGTCTGACAGTACCGGCGCATCTCTGACTACCAAGATCCATGAGCTTAAACAGGCC 1440
Db 1381 ACGGTCTGACAGTACCGGCGCATCTCTGACTACCAAGATCCATGAGCTTAAACAGGCC 1440
QY 1441 AAAGTGACACAGTACGATACCGTTCGCGGAGGCTTCGGGACTCCCTGGGCGGAAT 1500
Db 1441 AAAGTGACACAGTACGATACCGTTCGCGGAGGCTTCGGGACTCCCTGGGCGGAAT 1500
QY 1501 GTGGACCTGGGCTGACAGGAACCATGGAGTTGTGCACCCCACTCTTGGGAGAGTGC 1560
Db 1501 GTGGACCTGGGCTGACAGGAACCATGGAGTTGTGCACCCCACTCTTGGGAGAGTGC 1560
QY 1561 AACAAAGTTTGAAGGAGAACACCTACTGCTTCGGCGGACCGGTGGAGACACTGCTCCGC 1620
Db 1561 AACAAAGTTTGAAGGAGAACACCTACTGCTTCGGCGGACCGGTGGAGACACTGCTCCGC 1620
QY 1621 TTTGGCAAGACCATCATGGAGGACAGCTGTGCTCGGGGCCAGCGCTCCATCCGATTGGGCTC 1680
Db 1621 TTTGGCAAGACCATCATGGAGGACAGCTGTGCTCGGGGCCAGCGCTCCATCCGATTGGGCTC 1680
QY 1681 AACCTGTATGGCATGACGGCGGTGCTGCTCGGGGCCAGCGCTCCATCCGATTGGGCTC 1740
Db 1681 AACCTGTATGGCATGACGGCGGTGCTGCTCGGGGCCAGCGCTCCATCCGATTGGGCTC 1740
QY 1741 CGCAACACACACAGGAGTCTCTTGGCCCAACACTTCTGCGTGGAGGTTACTTGCAG 1800
Db 1741 CGCAACACACAGGAGTCTCTTGGCCCAACACTTCTGCGTGGAGGTTACTTGCAG 1800
QY 1801 AATCTTTCAGCCCTCTCTAGCTGGCAAGTATGCTCCAGAAACCTAGATGACAGATT 1860
Db 1801 AATCTTTCAGCCCTCTCTAGCTGGCAAGTATGCTCCAGAAACCTAGATGACAGATT 1860
QY 1861 AAGAAAGTCCAGCAGATCTTGAGAGCGAGCCTATATCTGTGCCCAACCTCTGGAC 1920
Db 1861 AAGAAAGTCCAGCAGATCTTGAGAGCGAGCCTATATCTGTGCCCAACCTCTGGAC 1920
QY 1921 AGGACATGCTGAGGAGGAGACAGTGTCCCTGCTACCGCCCGCCCTACCCATGCCCG 1980
Db 1921 AGGACATGCTGAGGAGGAGACAGTGTCCCTGCTACCGCCCGCCCTACCCATGCCCG 1980
QY 1981 TTGCTGGATGACTGTTACTCTTTTTCAGAGGTGTTGGATATACAGGTTAAGCCTT 2040
Db 1981 TTGCTGGATGACTGTTACTCTTTTTCAGAGGTGTTGGATATACAGGTTAAGCCTT 2040
QY 2041 TTGTTCCCGTCTGCACCTGAAGGTTGTGCGCTGCGCTGGGAGAGCCTCTTCCAGGTT 2100
Db 2041 TTGTTCCCGTCTGCACCTGAAGGTTGTGCGCTGCGCTGGGAGAGCCTCTTCCAGGTT 2100
QY 2101 TGACCTGACGAGTCTCTTAACAGGACCATCATGAGTCTGAACTGAGCGGAGAGA 2160
Db 2101 TGACCTGACGAGTCTCTTAACAGGACCATCATGAGTCTGAACTGAGCGGAGAGA 2160
QY 2161 GAGATGGAATGCTGACCCCTGGAAGTGTGCGCTGCGCTGGGAGAGCCTCTTGAAGCTCT 2220
Db 2161 GAGATGGAATGCTGACCCCTGGAAGTGTGCGCTGCGCTGGGAGAGCCTCTTGAAGCTCT 2220
QY 2221 AGTGAACATGGGGCTTATGCTGCTCCCTCAGGAGTGTGAGTGGGTGGGACCTGTGTC 2280
Db 2221 AGTGAACATGGGGCTTATGCTGCTCCCTCAGGAGTGTGAGTGGGTGGGACCTGTGTC 2280
QY 2281 AGGTGTGATAGCCATTTCTGCTCAACACACATTTCTTAAGAACAGCTTGAAGCTCT 2340
Db 2281 AGGTGTGATAGCCATTTCTGCTCAACACACATTTCTTAAGAACAGCTTGAAGCTCT 2340
QY 2341 GTCTGGGTCAATTCATTTAACTAGAGGAGGACATTTAAACATGTACACAGGACCAT 2400
Db 2341 GTCTGGGTCAATTCATTTAACTAGAGGAGGACATTTAAACATGTACACAGGACCAT 2400
```

```
QY 2401 TAACAAGAATATAAAATGTACAATCTGTGTACTCTTAAAAAATAAAAAA 2452
Db 2401 TAACAAGAATATAAAATGTACAATCTGTGTACTCTTAAAAAATAAAAAA 2452

RESULT 2
ABV211123
ID ABV211123 standard; cDNA; 2522 BP.
XX AC ABV211123;
XX AC ABV211123;
XX DT 13-SEP-2002 (first entry)
XX XX Human prostate expression marker cDNA 21114.
XX DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US05171.
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-189862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-219007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PA Schlegel R, Endege WO, Monahan JE;
XX PI WPI; 2001-662795/76.
XX DR Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 3496; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV52213) of the
XX CC specification or its complement. (I) is useful for:
XX CC (a) assessing whether a patient is afflicted with prostate cancer;
XX CC (b) monitoring the progression of prostate cancer in a patient;
XX CC (c) assessing the efficacy of a test compound to inhibit prostate
XX CC cancer in a patient;
XX CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX CC in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) assessing the prostate cell carcinogenic potential of a compound;
XX CC (g) determining whether prostate cancer has metastasized in a patient;
XX CC (h) assessing the aggressiveness or indolence of prostate cancer in a
XX CC patient;
XX CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX SQ Sequence 2522 BP; 642 A; 634 C; 701 G; 543 T; 2 other;

Query Match 100.0%; Score 2452; DB 23; Length 2522;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTGTGTCTCCCTCGCGCTAAGAGGGAGACTGAGGCTGAGGCTGGGGAACATCG 60
Db 51 CGTGTGTGTCTCCCTCGCGCTAAGAGGGAGACTGAGGCTGAGGCTGGGGAACATCG 110
QY 61 GGCAGCATGAGCGGCTGCGGGCTCTCTCCGCGCACACAGGCTGCGGCTGCTGCCGG 120
Db 61 GGCAGCATGAGCGGCTGCGGGCTCTCTCCGCGCACACAGGCTGCGGCTGCTGCCGG 120
```


Db	111	GGCAGCATGAGCGGCTGCGGGCTCTTCCTGCGCACCACGGCTCGGCTGCTGCGCGG	170
Qy	121	GGTCTGGTGTCTCTACCGCGAACCGCGGGCTACTGCGCACCGCCCGCTGTAGAGCT	180
Db	171	GGTCTGGTGTCTCTACCGCGAACCGCGGGCTACTGCGCACCGCCCGCTGTAGAGCT	230
Qy	181	TTGCGCCAAAGAGCTTTTCTAGCAAAATCAAGAAAGAAAGTTTCCCATTTCCAGAA	240
Db	231	TTGCGCCAAAGAGCTTTTCTAGCAAAATCAAGAAAGAAAGTTTCCCATTTCCAGAA	290
Qy	241	GTTAGCAAGATGAACCTTAATGAATCAATCACTTCTGGGACCCCGTGAAAAATTTCTC	300
Db	291	GTTAGCAAGATGAACCTTAATGAATCAATCACTTCTGGGACCCCGTGAAAAATTTCTC	350
Qy	301	ACTGAAGAGTGGACTCCCGAAATAATTGACCAAGGAAGGAATCCAGATGAACCTTTC	360
Db	351	ACTGAAGAGTGGACTCCCGAAATAATTGACCAAGGAAGGAATCCAGATGAACCTTTC	410
Qy	361	GAGAAATGAAGAGCTTAGGCTTTTGGGCTGCAAGTCCAGAGAAATATGCTGGCCTG	420
Db	411	GAGAAATGAAGAGCTTAGGCTTTTGGGCTGCAAGTCCAGAGAAATATGCTGGCCTG	470
Qy	421	GGCTTCTCCAAACCATCTACTCAAGACTAGGGGAGATCATAGCATGGATGGGTCATC	480
Db	471	GGCTTCTCCAAACCATCTACTCAAGACTAGGGGAGATCATAGCATGGATGGGTCATC	530
Qy	481	ACTGTGACCTGGCAGCGCACCAGGCTATTGGCTCAAGGGGATCATCTTTGGCTGGCACT	540
Db	531	ACTGTGACCTGGCAGCGCACCAGGCTATTGGCTCAAGGGGATCATCTTTGGCTGGCACT	590
Qy	541	GAGGAGCAAGAACCAATACTTGCCTAAATGCGCTCCGGGGAGCAGATTTGAGCGCTTC	600
Db	591	GAGGAGCAAGAACCAATACTTGCCTAAATGCGCTCCGGGGAGCAGATTTGAGCGCTTC	650
Qy	601	TGCCTCAGGAGCAGCCAGTGGGAGGATCCAGCTCAATCCGAGCAGAGCCACACTA	660
Db	651	TGCCTCAGGAGCAGCCAGTGGGAGGATCCAGCTCAATCCGAGCAGAGCCACACTA	710
Qy	661	AGTGAAGACAAGAACCTACATCCTCAATGGCTCCAGGCTCTGGATTAATTTGAGGA	720
Db	711	AGTGAAGACAAGAACCTACATCCTCAATGGCTCCAGGCTCTGGATTAATTTGAGGA	770
Qy	721	CTGGCCAAATTTTACTGTGTTGCCAAGACTGAGGTCTGTGATTTCTGATGGATCAGTG	780
Db	771	CTGGCCAAATTTTACTGTGTTGCCAAGACTGAGGTCTGTGATTTCTGATGGATCAGTG	830
Qy	781	AAAGACAATAACACAGCATTCATAGTAGAAAGAGACTTTGGTGGAGTCACTTAATTTGGGAAA	840
Db	831	AAAGACAATAACACAGCATTCATAGTAGAAAGAGACTTTGGTGGAGTCACTTAATTTGGGAAA	890
Qy	841	CCCGAAGATAAATTAGGCATTCGGGGCTCCAAACACTTGTGAAGTCCATTTTGAAGAACCC	900
Db	891	CCCGAAGATAAATTAGGCATTCGGGGCTCCAAACACTTGTGAAGTCCATTTTGAAGAACCC	950
Qy	901	AGATACCTGTGGAAACATCCTTGAGAGTTCGAGATGGGTTTAAAGTGGCCATGAAC	960
Db	951	AGATACCTGTGGAAACATCCTTGAGAGTTCGAGATGGGTTTAAAGTGGCCATGAAC	1010
Qy	961	ATCCTCAACAGCGCGGTTACGATGGGACGCTCGTGGCTGGCTGCTCAAGAGATTG	1020
Db	1011	ATCCTCAACAGCGCGGTTACGATGGGACGCTCGTGGCTGGCTGCTCAAGAGATTG	1070
Qy	1021	ATTGAATGACTGCTGAGTACCGCTCCAGAGAACAGTTTAAAGAGGCTCAGTGA	1080
Db	1071	ATTGAATGACTGCTGAGTACCGCTCCAGAGAACAGTTTAAAGAGGCTCAGTGA	1130
Qy	1081	TTTGGATTGATTGAGGAGAAATTTGACATGATGGCTCAGAGGCTTAAGTCAATGAGAGT	1140
Db	1131	TTTGGATTGATTGAGGAGAAATTTGACATGATGGCTCAGAGGCTTAAGTCAATGAGAGT	1190
Qy	1141	ATGACCTACCTACAGCAGGATGCTGGACCAACCTGGCTTTCCGACTGCTCCATCGAG	1200
Db	1191	ATGACCTACCTACAGCAGGATGCTGGACCAACCTGGCTTTCCGACTGCTCCATCGAG	1250
Qy	1201	GCAGCCATGTTGAAGTGTTCAGCTCCGAGGCGCCCTGGCAGTGTGTGAGTGGAGCGCTG	1260
Db	1251	GCAGCCATGTTGAAGTGTTCAGCTCCGAGGCGCCCTGGCAGTGTGTGAGTGGAGCGCTG	1310
Qy	1261	CAGATCTCGGGGGCTTGGGCTACACAAGGAGACTATCCGTACGAGCGCATCTGCGTGAC	1320
Db	1311	CAGATCTCGGGGGCTTGGGCTACACAAGGAGACTATCCGTACGAGCGCATCTGCGTGAC	1370
Qy	1321	ACCGCATCTCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTATACATCGCCCTG	1380
Db	1371	ACCGCATCTCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTATACATCGCCCTG	1430
Qy	1381	ACGGGTCTGACAGCATGCGCGCGCATCTCTGACTACCAGGATCCATGAGCTTTAAACAGGCC	1440
Db	1431	ACGGGTCTGACAGCATGCGCGCGCATCTCTGACTACCAGGATCCATGAGCTTTAAACAGGCC	1490
Qy	1441	AAAGTGACACAGTCAATGATACCGTTGGCCGGAGCTTCGGGACTCCCTGGGCGGAAC	1500
Db	1491	AAAGTGACACAGTCAATGATACCGTTGGCCGGAGCTTCGGGACTCCCTGGGCGGAAC	1550
Qy	1501	GTGGACCTGGGGCTGACAGGCAACCATGGAGTTGTGACCCCGAGTCTTTCGCGAGAGTGCC	1560
Db	1551	GTGGACCTGGGGCTGACAGGCAACCATGGAGTTGTGACCCCGAGTCTTTCGCGAGAGTGCC	1610
Qy	1561	AAACAAGTTTGAGGAGAACACACTACTGCTCGGCGGACCGTGGAGACACTGCTGCTCGGC	1620
Db	1611	AAACAAGTTTGAGGAGAACACACTACTGCTCGGCGGACCGTGGAGACACTGCTGCTCGGC	1670
Qy	1621	TTTGGCAAGACCATCATGGAGGAGCAGCTGCTGCTGCGCAACACCTTCTGCGTGGAGCTTCTG	1680
Db	1671	TTTGGCAAGACCATCATGGAGGAGCAGCTGCTGCTGCGCAACACCTTCTGCGTGGAGCTTCTG	1730
Qy	1681	AACCTGTATGGATGATGAGCGCGTGTGCTCGCGGCGACCGCTCCATCGCAATTTGGGCTC	1740
Db	1731	AACCTGTATGGATGATGAGCGCGTGTGCTCGCGGCGACCGCTCCATCGCAATTTGGGCTC	1790
Qy	1741	CGCAACACGACCCACAGAGTCTCTTCTGCGCAACACCTTCTGCGTGGAGCTTCTG	1800
Db	1791	CGCAACACGACCCACAGAGTCTCTTCTGCGCAACACCTTCTGCGTGGAGCTTCTG	1850
Qy	1801	AATCTCTTCAGCCCTCTCTCAGCTGGACAAAGTATGCTCCAGAAACCTAGATGAGCAGATT	1860
Db	1851	AATCTCTTCAGCCCTCTCTCAGCTGGACAAAGTATGCTCCAGAAACCTAGATGAGCAGATT	1910
Qy	1861	AAGAAAGTGTCCAGCAGATCCCTTGTGAGAGCGAGCCTATATCTGTGCCACCCCTCTGGAC	1920
Db	1911	AAGAAAGTGTCCAGCAGATCCCTTGTGAGAGCGAGCCTATATCTGTGCCACCCCTCTGGAC	1970
Qy	1921	AGGACATGTGAGCGAGGGGACAGTGTCCCTCTACCGCCCGCCCTACCCATGGCCCG	1980
Db	1971	AGGACATGTGAGCGAGGGGACAGTGTCCCTCTACCGCCCGCCCTACCCATGGCCCG	2030
Qy	1981	TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2040
Db	2031	TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2090
Qy	2041	TTGTTTCCCGCTCTGCACTGAAAGGTTGTGCGCTGGGCTGGGAGAGGCTTCTCCAGGTTT	2100
Db	2091	TTGTTTCCCGCTCTGCACTGAAAGGTTGTGCGCTGGGCTGGGAGAGGCTTCTCCAGGTTT	2150
Qy	2101	TGACCTCGAGCGAGTGTCTCTACAGGACCATCACAGCTTCTGAACTGAGCGCGAGAGA	2160
Db	2151	TGACCTCGAGCGAGTGTCTCTACAGGACCATCACAGCTTCTGAACTGAGCGCGAGAGA	2210
Qy	2161	GAGAAATGAAATGCTGACCCCTGGAAGTGGCGGATTTCTGGTCAATTTGAGGAGACCAT	2220
Db	2211	GAGAAATGAAATGCTGACCCCTGGAAGTGGCGGATTTCTGGTCAATTTGAGGAGACCAT	2270
Qy	2221	AGTGAAGACTGGGCTTATGCTGCTCCAGGCTGAGGTGGGTGGGAGCTGTGTC	2280
Db	2271	AGTGAAGACTGGGCTTATGCTGCTCCAGGCTGAGGTGGGTGGGAGCTGTGTC	2330

QY 2281 AGGTGGATAGCCATTTCCTCAACACACACATTCCTTAAGAAACAGCTTGAAGCTCT 2340
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2331 AGGTGGATAGCCATTTCCTCAACACACACATTCCTTAAGAAACAGCTTGAAGCTCT 2390
 QY 2341 GTCTGGGTCAATTTAAACTAGACGAGGACACTTAAACATGTACCAGGAACCAATT 2400
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2391 GTCTGGGTCAATTTAAACTAGACGAGGACACTTAAACATGTACCAGGAACCAATT 2450
 QY 2401 TAACAAAGATATATAAATGTCACAATCTGTACTGTATAAAAAAAAAAAAAA 2452
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 2451 TAACAAAGATATATAAATGTCACAATCTGTACTGTATAAAAAAAAAAAAAA 2502

RESULT 3

ABV23498
 ID ABV23498 standard; cDNA; 2522 BP.

XX
 AC ABV23498;

DT 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 23489.

DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.

XX Homo sapiens.

OS WO20010860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US05171.

XX 17-FEB-2000; 2000US-183319P.

PR 16-MAR-2000; 2000US-189862P.

PR 25-MAY-2000; 2000US-207454P.

PR 09-JUN-2000; 2000US-211314P.

PR 18-JUL-2000; 2000US-219007P.

PR 13-DEC-2000; 2000US-255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -

PS Claim 1; Page 4299-4300; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;

CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;

CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 2522 BP; 642 A; 634 C; 701 G; 543 T; 2 other;

Query Match 100.0%; Score 2452; DB 23; Length 2522;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCGTGTGTGTCCTCGCGCGCTAAGAGGGGAGAGCTGAGGCTGAGGCTGGGGAACATCG 60
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 51 CGTGTGTGTCCTCGCGCGCTAAGAGGGGAGAGCTGAGGCTGAGGCTGGGGAACATCG 110
 QY 61 GCGAGCATGAGCGGTGCGGGCTCTTCTGCGCACACGCGCTGCGGCTGCGCTGCCGG 120
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 111 GCGAGCATGAGCGGTGCGGGCTCTTCTGCGCACACGCGCTGCGGCTGCGCTGCCGG 170
 QY 121 GGTCTGCGGTCTCTACCGCGAACCAGCGGCTACTGCGCACACGCGCGCTGTACGAGCT 180
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 171 GGTCTGCGGTCTCTACCGCGAACCAGCGGCTACTGCGCACACGCGCGCTGTACGAGCT 230
 QY 181 TTGCGCAAGAGCTTTTCTAGGCAAAATCAAGAGAAAGAAAGTTCCTCCCATTTCCAGAA 240
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 231 TTGCGCAAGAGCTTTTCTAGGCAAAATCAAGAGAAAGAAAGTTCCTCCCATTTCCAGAA 290
 QY 241 GTTAGCCAAAGATGAACCTTAATGAATCAATCAGTTCCTTGGGACCGCTGGAAAAATCTTC 300
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 291 GTTAGCCAAAGATGAACCTTAATGAATCAATCAGTTCCTTGGGACCGCTGGAAAAATCTTC 350
 QY 301 ACTGAAGAGGTGGACTCCCGAAAAATTGACGAGGAGGAAATCCAGATGAACCTTTG 360
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 351 ACTGAAGAGGTGGACTCCCGAAAAATTGACGAGGAGGAAATCCAGATGAACCTTTG 410
 QY 361 GAGAAATTGAAGAGCTTAGGCTTTTGGGCTGCAAGTCCAGAGGAGATCATGAGCATGGTGCATC 420
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 411 GAGAAATTGAAGAGCTTAGGCTTTTGGGCTGCAAGTCCAGAGGAGATCATGAGCATGGTGCATC 470
 QY 421 GGCTTCTCCACACCATGTACTCAAGACTAGGGGAGATCATGAGCATGGTGCATC 480
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 471 GGCTTCTCCACACCATGTACTCAAGACTAGGGGAGATCATGAGCATGGTGCATC 530
 QY 481 ACTGTACCCCTGGCAGCGCACAGGCTATTGGGCTCAAGGGGATCATCTTGGCTGGCACT 540
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 531 ACTGTACCCCTGGCAGCGCACAGGCTATTGGGCTCAAGGGGATCATCTTGGCTGGCACT 590
 QY 541 GAGGAGCAGAAAGCCAAATACTTGCCTAAACTGGCTCCGGGAGGACACATTTGAGGCTTC 600
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 591 GAGGAGCAGAAAGCCAAATACTTGCCTAAACTGGCTCCGGGAGGACACATTTGAGGCTTC 650
 QY 601 TGCCCTACGGAGCGCAGCCAGTGGGAGGATGACGCTCAATCCGAGGAGAGCCACACTA 660
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 651 TGCCCTACGGAGCGCAGCCAGTGGGAGGATGACGCTCAATCCGAGGAGAGCCACACTA 710
 QY 661 AGTGAAGACAAAGAGCACTACATCCCAATGGCTCCCAAGGCTCTGGATTACTAATGGAGGA 720
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 711 AGTGAAGACAAAGAGCACTACATCCCAATGGCTCCCAAGGCTCTGGATTACTAATGGAGGA 770
 QY 721 CTGCGCAATATTTTACTCTGTGTTGCAAGAGCTGAGGCTGTTGATTCTGATGATGATG 780
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 771 CTGCGCAATATTTTACTCTGTGTTGCAAGAGCTGAGGCTGTTGATTCTGATGATGATG 830
 QY 781 AAAGACAAATACACGATTCATAGTAGAGAGAGACTTTTGGTGGAGTCACTAATGGGAAA 840
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 831 AAAGACAAATACACGATTCATAGTAGAGAGAGACTTTTGGTGGAGTCACTAATGGGAAA 890
 QY 841 CCGGAGATTAATTAGGCATTCGGGCTCCACACTTTGGAAGTCCATTTTGAAGAACACC 900
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 891 CCGGAGATTAATTAGGCATTCGGGCTCCACACTTTGGAAGTCCATTTTGAAGAACACC 950
 QY 901 AAGATACCTGTGGAAAAACATCTTGGAGAGGTGCGGAGATGGGTTTANGGTGCCCATGAAC 960
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 951 AAGATACCTGTGGAAAAACATCTTGGAGAGGTGCGGAGATGGGTTTANGGTGCCCATGAAC 1010
 QY 961 ATCCTCAACAGCGCGGTTTCAGCATGGGACGCTGCTGGCTGGGCTGCTCAAGAGATTG 1020
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 1011 ATCCTCAACAGCGCGGTTTCAGCATGGGACGCTGCTGGCTGGGCTGCTCAAGAGATTG 1070
 QY 1021 ATTGAATACCTGCTGAGTACCGCTGCACAGGAAACAGTTTAAACAGAGGCTCAGTGA 1080
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 1071 ATTGAATACCTGCTGAGTACCGCTGCACAGGAAACAGTTTAAACAGAGGCTCAGTGA 1130

QY	1081	TTTGGATTGATTCAGGAGAAATTTGACATGATGGCTCAGAGGCTTACGTCATGAGAGT	1140
Db	1131	TTTGGATTGATTCAGGAGAAATTTGACATGATGGCTCAGAGGCTTACGTCATGAGAGT	1190
QY	1141	ATGACCTACTCACAGCAGGATGCTGGACCAACTGGCTTTCCGAGCTGCTCCATCGAG	1200
Db	1191	ATGACCTACTCACAGCAGGATGCTGGACCAACTGGCTTTCCGAGCTGCTCCATCGAG	1250
QY	1201	GCAGCCATGGTGAAGGTGTTACGCTCCGAGGCGCCTGGCAGTGTGTGAGTGAGCGGTG	1260
Db	1251	GCAGCCATGGTGAAGGTGTTACGCTCCGAGGCGCCTGGCAGTGTGTGAGTGAGCGGTG	1310
QY	1261	CAGATCCTCGGGGCTTGGGCTACACAGGGGACTATCCGTACAGGCACTACGCTGAC	1320
Db	1311	CAGATCCTCGGGGCTTGGGCTACACAGGGGACTATCCGTACAGGCACTACGCTGAC	1370
QY	1321	ACCCGATCCTCCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCCCCCTG	1380
Db	1371	ACCCGATCCTCCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCCCCCTG	1430
QY	1381	ACGGGTCTGAGCATCGCGCCGCATCTTGCTGCGGAGGCTTCCGGACTCCCTGGGCGGAAT	1440
Db	1431	ACGGGTCTGAGCATCGCGCCGCATCTTGCTGCGGAGGCTTCCGGACTCCCTGGGCGGAAT	1490
QY	1441	AAAGTGAGCACAGTATGATACCTTTGCGGAGGCTTCCGGACTCCCTGGGCGGAAT	1500
Db	1491	AAAGTGAGCACAGTATGATACCTTTGCGGAGGCTTCCGGACTCCCTGGGCGGAAT	1550
QY	1501	GTGACCTGGGGCTGACAGCAACCATGAGATTGTGACCCAGTCTTGGGACAGTGC	1560
Db	1551	GTGACCTGGGGCTGACAGCAACCATGAGATTGTGACCCAGTCTTGGGACAGTGC	1610
QY	1561	AACAAGTTGAGGAGACACTACTGCTCGCGGAGGCTGAGACACTGCTGCTCGC	1620
Db	1611	AACAAGTTGAGGAGACACTACTGCTCGCGGAGGCTGAGACACTGCTGCTCGC	1670
QY	1621	TTTGGCAAGCAATCATGAGGAGCAGCTGCTACTGAAGCGGTGGCCAAACATCCTCATC	1680
Db	1671	TTTGGCAAGCAATCATGAGGAGCAGCTGCTACTGAAGCGGTGGCCAAACATCCTCATC	1730
QY	1681	AACCTGTATGGCATGCGGCGCTGCTGCGGCGGAGCGCTCCATCCGATTTGGGCTC	1740
Db	1731	AACCTGTATGGCATGCGGCGCTGCTGCGGCGGAGCGCTCCATCCGATTTGGGCTC	1790
QY	1741	CGCAACCAACGACGAGGTTCTCTGGCCCAACACCTTCTGCGTGAAGCTTACTTGCAG	1800
Db	1791	CGCAACCAACGACGAGGTTCTCTGGCCCAACACCTTCTGCGTGAAGCTTACTTGCAG	1850
QY	1801	AATCTCTTACGCTCTCAGCTGGACAAGTATGCTCCAGAAACCTAGATGAGCAGATT	1860
Db	1851	AATCTCTTACGCTCTCAGCTGGACAAGTATGCTCCAGAAACCTAGATGAGCAGATT	1910
QY	1861	AAGAAGTGTCGCCAGCAGATCCTTTGAGAAGCGGCTTATCTGTGCCCCACCTCTGGAC	1920
Db	1911	AAGAAGTGTCGCCAGCAGATCCTTTGAGAAGCGGCTTATCTGTGCCCCACCTCTGGAC	1970
QY	1921	AGGACATGCTGAGGAGGGGACAGTGTCCCTGTACGCCCGCCCTACCCATGGCCCG	1980
Db	1971	AGGACATGCTGAGGAGGGGACAGTGTCCCTGTACGCCCGCCCTACCCATGGCCCG	2030
QY	1981	TGCTGGATGACTGTACTCTTTTTCAGAAGGTGTTGGGATTATCAGAGTTAAGCCTT	2040
Db	2031	TGCTGGATGACTGTACTCTTTTTCAGAAGGTGTTGGGATTATCAGAGTTAAGCCTT	2090
QY	2041	TTTGTCCCGCTGACCATGAAGGTTGTCGCCCTGGGCTGGGAGGCTCTTCCAGGTTT	2100
Db	2091	TTTGTCCCGCTGACCATGAAGGTTGTCGCCCTGGGCTGGGAGGCTCTTCCAGGTTT	2150
QY	2101	TCACCTGAGGAGGCTCTCTACAGGACCATCACAGCTTCTGAACCTGAGCCCGGAGA	2160
Db	2151	TCACCTGAGGAGGCTCTCTACAGGACCATCACAGCTTCTGAACCTGAGCCCGGAGA	2210

QY	2161	GAGAAATGAATGCTGACCCCTGGAAGTGGGGGTATTTGTCATTTGAGGAGACCAT	2220
Db	2211	GAGAAATGAATGCTGACCCCTGGAAGTGGGGGTATTTGTCATTTGAGGAGACCAT	2270
QY	2221	AGTGAACACTGGGCTTATGCTGCTGCTCCAGGCTGAGGTGGGTGGGACCTGTGTC	2280
Db	2271	AGTGAACACTGGGCTTATGCTGCTGCTCCAGGCTGAGGTGGGTGGGACCTGTGTC	2330
QY	2281	AGTGTGTGATAGCCATTTCTGCTCAACACACACATTTCTTAAGAAACAGCTTGAAGCTCT	2340
Db	2331	AGTGTGTGATAGCCATTTCTGCTCAACACACACATTTCTTAAGAAACAGCTTGAAGCTCT	2390
QY	2341	GTCTGGGTTCATTTAACTAGAGGAGGACCTTAAACATGTACCAGGAACCATT	2400
Db	2391	GTCTGGGTTCATTTAACTAGAGGAGGACCTTAAACATGTACCAGGAACCATT	2450
QY	2401	TAAACAAGATATAAATGTCAATCTGTCTACTGTATAAAAAA	2452
Db	2451	TAAACAAGATATAAATGTCAATCTGTCTACTGTATAAAAAA	2502

RESULT 4

ABV26965

ID ABV26965 standard; cDNA; 2522 BP.

XX ABV26965;

AC ABV26965;

DT 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 26956.

DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US05171.

XX 17-FEB-2000; 2000US-183319P.

PR 16-MAR-2000; 2000US-189862P.

PR 25-MAY-2000; 2000US-207454P.

PR 09-JUN-2000; 2000US-211314P.

PR 18-JUL-2000; 2000US-219007P.

PR 13-DEC-2000; 2000US-255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer

XX Claim 1; Page 5464; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (1) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (1) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;

CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound;

CC (g) determining whether prostate cancer has metastasized in a patient;

CC	(h) assessing the aggressiveness or indolence of prostate cancer in a patient;									
CC	(i) is also useful as a pharmacodynamic or pharmacogenomic marker.									
XX										
SQ	Sequence 2522 BP; 642 A; 634 C; 701 G; 543 T; 2 other;									
	Query Match									
	Best Local Similarity 100.0%; Score 2452; DB 23; Length 2522;									
	Matches 2452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	CCTGTGTGTGTCCTCCGCGCTAGAGAGGAGACTGAGGCTGAGGCTGGGGAACATCG	60							
DB	51	CCTGTGTGTGTCCTCCGCGCTAGAGAGGAGACTGAGGCTGAGGCTGGGGAACATCG	110							
QY	61	GGCAGCATGAGCGGTGGGCTCTCTTCCGCGCACCAGCGGTGGGCTGCTGCTGCGCG	120							
DB	111	GGCAGCATGAGCGGTGGGCTCTCTTCCGCGCACCAGCGGTGGGCTGCTGCTGCGCG	170							
QY	121	GCTCTGCGTCTCTACCGCAACCGGGGCTACTGCGCACCGCCGCTGTACGAGCT	180							
DB	171	GCTCTGCGTCTCTACCGCAACCGGGGCTACTGCGCACCGCCGCTGTACGAGCT	230							
QY	181	TTGCCAAAGAGCTTTTCTAGGCAAAATCAAGAGAAAGAAATTTCCCATTTCCAGAA	240							
DB	231	TTGCCAAAGAGCTTTTCTAGGCAAAATCAAGAGAAAGAAATTTTCCCATTTCCAGAA	290							
QY	241	GTTAGCCAAAGATGAACCTTAATGAATCAATCAGTTCTTGGGACCGGTGGAAATTTCTC	300							
DB	291	GTTAGCCAAAGATGAACCTTAATGAATCAATCAGTTCTTGGGACCGGTGGAAATTTCTC	350							
QY	301	ACTCAAGAGTGGACTCCCAAAATTTGACCAAGAGGAAATTTCCAGATGAACCTTTG	360							
DB	351	ACTCAAGAGTGGACTCCCAAAATTTGACCAAGAGGAAATTTTCCAGATGAACCTTTG	410							
QY	361	GAGAAATTTGAGAGCTTAGGCTTTTGGGCTGCAAGTCCCAAGAAATATGTCGCTG	420							
DB	411	GAGAAATTTGAGAGCTTAGGCTTTTGGGCTGCAAGTCCCAAGAAATATGTCGCTG	470							
QY	421	GGCTTCTCCCAACACCACTGACTCAAGACTAGGGAGATCATCAGCTGGTCCATC	480							
DB	471	GGCTTCTCCCAACACCACTGACTCAAGACTAGGGAGATCATCAGCTGGTCCATC	530							
QY	481	ACTGTGACCTGGCAGCGCACCAAGCTATTGGCTCAAGGAGATCATCTGGCTGGCCT	540							
DB	531	ACTGTGACCTGGCAGCGCACCAAGCTATTGGCTCAAGGAGATCATCTGGCTGGCCT	590							
QY	541	GAGGAGCAAAACCAATCTTGTCTAACTGGGCTCCGGGAGCACATTTGACGCTTC	600							
DB	591	GAGGAGCAAAACCAATCTTGTCTAACTGGGCTCCGGGAGCACATTTGACGCTTC	650							
QY	601	TGCTTCAGGAGCGCACCAAGTGGGAGCTGAGCTCAATCCGAGAGAGCCACATA	660							
DB	651	TGCTTCAGGAGCGCACCAAGTGGGAGCTGAGCTCAATCCGAGAGAGCCACATA	710							
QY	661	AGTGAAGACAAGAGCACTACATCTCAATGGCTCCAAAGCTGAGTACTAATGGAGGA	720							
DB	711	AGTGAAGACAAGAGCACTACATCTCAATGGCTCCAAAGCTGAGTACTAATGGAGGA	770							
QY	721	CTGGCCAAATTTTACTGTGTTTCCAAAGACTGAGTCTGATTTGATGGATCAGTG	780							
DB	771	CTGGCCAAATTTTACTGTGTTTCCAAAGACTGAGTCTGATTTGATGGATCAGTG	830							
QY	781	AAAGACAAATCACAGCACTCATAGTAGAAGAGACTTTGTTGGAGTCACTAATGGGAA	840							
DB	831	AAAGACAAATCACAGCACTCATAGTAGAAGAGACTTTGTTGGAGTCACTAATGGGAA	890							
QY	841	CCCGAAGATAAATAGGATTTGGGGCTCCCAACTTTGTGAAGTCCATTTTGAACACAC	900							
DB	891	CCCGAAGATAAATAGGATTTGGGGCTCCCAACTTTGTGAAGTCCATTTTGAACACAC	950							
QY	901	AAGATACCTGTGGAACATCTTTGGAGAGTCCGAGATGGTTTAAGTTAGGCTGAAC	960							
DB	951	AAGATACCTGTGGAACATCTTTGGAGAGTCCGAGATGGTTTAAGTTAGGCTGAAC	1010							

QY	961	ATCTCAACAGCGCGGTTTCAGCATGGCGAGCGTCTGGCTGGGCTGCTCAAGAGATTG	1020							
DB	1011	ATCTCAACAGCGCGGTTTCAGCATGGCGAGCGTCTGGCTGGGCTGCTCAAGAGATTG	1070							
QY	1021	ATTGAATGACTGCTGAGTACGCTGCACAAGGAAACAGTTTAAACAAGAGCTCAAGTAA	1080							
DB	1071	ATTGAATGACTGCTGAGTACGCTGCACAAGGAAACAGTTTAAACAAGAGCTCAAGTAA	1130							
QY	1081	TTTGGATTTGATTTCAGGAGAAATTTGCACTGATGGCTCAGAAGGCTTACGTATGAGAGT	1140							
DB	1131	TTTGGATTTGATTTCAGGAGAAATTTGCACTGATGGCTCAGAAGGCTTACGTATGAGAGT	1190							
QY	1141	ATGACCTACCTCACAGCAGGATGCTGGAACCAACCTGGCTTCCCGACTGCTCCATCCAG	1200							
DB	1191	ATGACCTACCTCACAGCAGGATGCTGGAACCAACCTGGCTTCCCGACTGCTCCATCCAG	1250							
QY	1201	GCAGCCATGGTCAAGGTGTTTCACTCCGAGCGGCTGGCAGTGTGTGAGTGAGGCGCTG	1260							
DB	1251	GCAGCCATGGTCAAGGTGTTTCACTCCGAGCGGCTGGCAGTGTGTGAGTGAGGCGCTG	1310							
QY	1261	CAGATCTCTCGGGGCTTGGGCTTACACAAGGAACTATCCGTACGAGCGCATACTGGTGAC	1320							
DB	1311	CAGATCTCTCGGGGCTTGGGCTTACACAAGGAACTATCCGTACGAGCGCATACTGGTGAC	1370							
QY	1321	ACCCGCATCTCTCTCATCTTCGAGGAACTATCCGTACGAGCGCATACTGGTGAC	1380							
DB	1371	ACCCGCATCTCTCTCATCTTCGAGGAACTATCCGTACGAGCGCATACTGGTGAC	1430							
QY	1381	ACGGCTCTGAGCATGCGCGGCTTGGGCTTACACAAGGAACTATCCGTACGAGCGCAT	1440							
DB	1431	ACGGCTCTGAGCATGCGCGGCTTGGGCTTACACAAGGAACTATCCGTACGAGCGCAT	1490							
QY	1441	AAAGTGACACAGTATGATGATGAGTGTGGCGGAGGCTTGGGAGTCTCCCTGGGCGGAACT	1500							
DB	1491	AAAGTGACACAGTATGATGATGAGTGTGGCGGAGGCTTGGGAGTCTCCCTGGGCGGAACT	1550							
QY	1501	GTGACCTGGGCTGACAGCAACCATGAGTGTGGTGTGACCCAGTCTTGGGAGAGTCC	1560							
DB	1551	GTGACCTGGGCTGACAGCAACCATGAGTGTGGTGTGACCCAGTCTTGGGAGAGTCC	1610							
QY	1561	AACAAGTTGAGGAGAACACCTACTGCTTGGCGGAGCGTGGAGACACTGCTGCTCCGC	1620							
DB	1611	AACAAGTTGAGGAGAACACCTACTGCTTGGCGGAGCGTGGAGACACTGCTGCTCCGC	1670							
QY	1621	TTTGGCAAGACCATCATGAGGAGCAGTGGTACTGAAGCGGTGGCCAACTCTCATC	1680							
DB	1671	TTTGGCAAGACCATCATGAGGAGCAGTGGTACTGAAGCGGTGGCCAACTCTCATC	1730							
QY	1681	AACCTGTATGGCATGAGCGCGTGTCTGCGGCGGAGCGCTCCATCCGCAATGGGCTC	1740							
DB	1731	AACCTGTATGGCATGAGCGCGTGTCTGCGGCGGAGCGCTCCATCCGCAATGGGCTC	1790							
QY	1741	CGCAACCAACCAAGGTTCTTCTTGGCCAACTCTTCTGCTGGAAGCTTACTTGGAG	1800							
DB	1791	CGCAACCAACCAAGGTTCTTCTTGGCCAACTCTTCTGCTGGAAGCTTACTTGGAG	1850							
QY	1801	AATCTCTTACGCTCTCTCAGCTGGACAGTATGCTCCAGAAACCTTAGTAGGAGATT	1860							
DB	1851	AATCTCTTACGCTCTCTCAGCTGGACAGTATGCTCCAGAAACCTTAGTAGGAGATT	1910							
QY	1861	AAGAAAGTGTCCAGCAGATCTCTTGGCCAACTATATCTGTGCCACCTCTGGAC	1920							
DB	1911	AAGAAAGTGTCCAGCAGATCTCTTGGCCAACTATATCTGTGCCACCTCTGGAC	1970							
QY	1921	AGGACATCTGAGGAGGAGCAGTGTCCCTGTACCGCGCGCCCTACCCATGGCCCG	1980							
DB	1971	AGGACATCTGAGGAGGAGCAGTGTCCCTGTACCGCGCGCCCTACCCATGGCCCG	2030							
QY	1981	TTGCTGGATGACTGTTACTCTTTTTCAGAAAGTGTGGGATTTACAGAGTTAAGCCTT	2040							
DB	2031	TTGCTGGATGACTGTTACTCTTTTTCAGAAAGTGTGGGATTTATCACAGGTTAAGCCTT	2090							

Qy	2041	TTGTTCCCTCTGCACTGAAGGTTGCGCTGCGCTGGGAGGCTCTCCAGGTTT	2100	CC	(b) monitoring the progression of prostate cancer in a patient;
Db	2091	TTGTTCCCTCTGCACTGAAGGTTGCGCTGCGCTGGGAGGCTCTCCAGGTTT	2150	CC	(c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;
Qy	2101	TGACCTCAGCGAGTGTCTTAACAGGACCATCACAGCTTCTGAAGCTGAGCCGGAGAGA	2160	CC	(d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;
Db	2151	TGACCTCAGCGAGTGTCTTAACAGGACCATCACAGCTTCTGAAGCTGAGCCGGAGAGA	2210	CC	(e) selecting a composition for inhibiting prostate cancer in a patient;
Qy	2161	GAGAAATGGAATGCTGACCCCTGGAACCTGGCGGGTATTCGTGTCATGAGGAGACCAAT	2220	CC	(f) assessing the prostate cell carcinogenic potential of a compound;
Db	2211	GAGAAATGGAATGCTGACCCCTGGAACCTGGCGGGTATTCGTGTCATGAGGAGACCAAT	2270	CC	(g) determining whether prostate cancer has metastasized in a patient;
Qy	2221	AGTGAACCTGGGCTTATCTGCTGCTCAGGGTGTAGGTGGTGGGACCTGTGTC	2280	CC	(h) assessing the aggressiveness or indolence of prostate cancer in a patient;
Db	2271	AGTGAACCTGGGCTTATCTGCTGCTCAGGGTGTAGGTGGTGGGACCTGTGTC	2330	CC	(I) is also useful as a pharmacodynamic or pharmacogenomic marker.
Qy	2281	AGGTGTGATAGCCATTCTTGAACACACACATCTCTAAGAAACAGCTTGAAGCTCT	2340	XX	Sequence 2522 BP; 642 A; 634 C; 701 G; 543 T; 2 other;
Db	2331	AGGTGTGATAGCCATTCTTGAACACACATCTCTAAGAAACAGCTTGAAGCTCT	2390	Qy	Query Match 100.0%; Score 2452; DB 23; Length 2522;
Qy	2341	GTCTGGGTCAATTTAACTAGAACGAGGACCTTAAACATGTACCAGAACCAIT	2400	Db	Best Local Similarity 100.0%; Pred. No. 0;
Db	2391	GTCTGGGTCAATTTAACTAGAACGAGGACCTTAAACATGTACCAGAACCAIT	2450	Qy	Matches 2452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	2401	TAACAAAGAAATATAAATGTACAAATCTGTACTGTAAATAAAAAA 2452		Db	1 CGTGTGTGTCTCCCTGCGGCTAAGAGGGGAGACTGAGGTGAGGTGGGAAACATCG 60
Db	2451	TAACAAAGAAATATAAATGTACAAATCTGTACTGTAAATAAAAAA 2502		Db	51 CGTGTGTGTCTCCCTGCGGCTAAGAGGGGAGACTGAGGTGAGGTGGGAAACATCG 110
RESULT 5				Qy	61 GGCAGCATGAGCGGCTGCGGCTCTTCTGCGCACACGGCTGCGGCTGCTGCCCTGCCG 120
ABV29355				Db	111 GGCAGCATGAGCGGCTGCGGCTCTTCTGCGCACACGGCTGCGGCTGCTGCCCTGCCG 170
XX	ABV29355 standard; cDNA; 2522 BP.			Qy	121 GGTCTGGTGTCTTACCGCGAACCGGGCTACTGCGCACCGCCGCTCTAGAGCT 180
AC				Db	171 GGTCTGGTGTCTTACCGCGAACCGGGCTACTGCGCACCGCCGCTCTAGAGCT 230
XX				Qy	181 TTGCGCAAGAGAGCTTTTCTAGGCAAAATCAAGAGAGAAAGATTTTCCATTTCCAGAA 240
XX	16-SEP-2002 (first entry)			Db	231 TTGCGCAAGAGAGCTTTTCTAGGCAAAATCAAGAGAGAAAGATTTTCCATTTCCAGAA 290
DE	Human prostate expression marker cDNA 29346.			Qy	241 GTTAGCCAAAGATGAACCTTAATGAATCAATCAGTCTTGGGACCCGCTGAAAAATCTTC 300
KW	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;			Db	291 GTTAGCCAAAGATGAACCTTAATGAATCAATCAGTCTTGGGACCCGCTGAAAAATCTTC 350
XX	pharmacogenomic marker; gene; ss.			Qy	301 ACTGAAGAGTGGACTCCCGAAATTTGACAGGAGGAGGAAATCCAGATGAAACTTTTG 360
OS	Homo sapiens.			Db	351 ACTGAAGAGTGGACTCCCGAAATTTGACAGGAGGAGGAAATCCAGATGAAACTTTTG 410
XX	WO200160860-A2.			Qy	361 GAGAAATGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGCCCTG 420
XX	23-AUG-2001.			Db	411 GAGAAATGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGCCCTG 470
PD				Qy	421 GGCTTCTCCAAACACCATGTACTCAAGACTAGGAGAGATCATCAGCATGGATGGTCCATC 480
XX	20-FEB-2001; 2001WO-US05171.			Db	471 GGCTTCTCCAAACACCATGTACTCAAGACTAGGAGAGATCATCAGCATGGATGGTCCATC 530
XX	17-FEB-2000; 2000US-183319P.			Qy	481 ACTGTGACCTGGCAGCGCACCGGCTTATGGGCTCAAGGGGATCATCTTGGCTGGCACT 540
PR	16-MAR-2000; 2000US-189862P.			Db	531 ACTGTGACCTGGCAGCGCACCGGCTTATGGGCTCAAGGGGATCATCTTGGCTGGCACT 590
PR	25-MAY-2000; 2000US-207454P.			Qy	541 GAGGAGCAGAAAGCAATATCTTCCCTAAACTGCGTCCGGGAGGACATTTGAGGCTTC 600
PR	09-JUN-2000; 2000US-211314P.			Db	591 GAGGAGCAGAAAGCAATATCTTCCCTAAACTGCGTCCGGGAGGACATTTGAGGCTTC 650
PR	18-JUL-2000; 2000US-219007P.			Qy	601 TGCCTCAGGAGCCAGCCAGTGGGAGGATGCGAGCTCAATCCGGAGCAGAGCCACACTA 660
PR	13-DEC-2000; 2000US-255281P.			Db	651 TGCCTCAGGAGCCAGCCAGTGGGAGGATGCGAGCTCAATCCGGAGCAGAGCCACACTA 710
XX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.			Qy	661 AGTGAAGACAAAGAGCACTACATCTCAATGGCTCCAGGTCTGGATTACTAATGGAGGA 720
XX	Schlegel R, Endege WO, Monahan JE;			Db	711 AGTGAAGACAAAGAGCACTACATCTCAATGGCTCCAGGTCTGGATTACTAATGGAGGA 770
XX	WPI; 2001-662795/76.			Qy	721 CTGCCCAATATTTTACTGTGTTGCAAGAGCTGAGGTCTGTGATTTCTGATGGATCAGTG 780
XX	Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -			Db	771 CTGCCCAATATTTTACTGTGTTGCAAGAGCTGAGGTCTGTGATTTCTGATGGATCAGTG 830
PT	Claim 1; Page 6287-6288; 11750pp; English.			Qy	781 AAAGACAAATACACAGCATTCATAGTAGAAAGAGAGCTTTGGTGGAGTCACTAATGGGAAA 840
PT				Db	831 AAAGACAAATACACAGCATTCATAGTAGAAAGAGAGCTTTGGTGGAGTCACTAATGGGAAA 890
CC	The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:				
CC	(a) assessing whether a patient is afflicted with prostate cancer;				

us-09-945-326-1.1.rng

Tue Sep 9 10:14:36 2003

PR	07-JUL-2000;	2000US-0216647.	PR	20-OCT-2000;	2000US-0241808.
PR	07-JUL-2000;	2000US-0216880.	PR	20-OCT-2000;	2000US-0241809.
PR	11-JUL-2000;	2000US-0217487.	PR	20-OCT-2000;	2000US-0241826.
PR	11-JUL-2000;	2000US-0217496.	PR	01-NOV-2000;	2000US-0244617.
PR	14-JUL-2000;	2000US-0218290.	PR	08-NOV-2000;	2000US-0246474.
PR	26-JUL-2000;	2000US-0220963.	PR	08-NOV-2000;	2000US-0246475.
PR	26-JUL-2000;	2000US-0220964.	PR	08-NOV-2000;	2000US-0246476.
PR	14-AUG-2000;	2000US-0224518.	PR	08-NOV-2000;	2000US-0246477.
PR	14-AUG-2000;	2000US-0224519.	PR	08-NOV-2000;	2000US-0246478.
PR	14-AUG-2000;	2000US-0225213.	PR	08-NOV-2000;	2000US-0246523.
PR	14-AUG-2000;	2000US-0225214.	PR	08-NOV-2000;	2000US-0246524.
PR	14-AUG-2000;	2000US-0225266.	PR	08-NOV-2000;	2000US-0246525.
PR	14-AUG-2000;	2000US-0225267.	PR	08-NOV-2000;	2000US-0246526.
PR	14-AUG-2000;	2000US-0225268.	PR	08-NOV-2000;	2000US-0246527.
PR	14-AUG-2000;	2000US-0225270.	PR	08-NOV-2000;	2000US-0246528.
PR	14-AUG-2000;	2000US-0225447.	PR	08-NOV-2000;	2000US-0246532.
PR	14-AUG-2000;	2000US-0225757.	PR	08-NOV-2000;	2000US-0246609.
PR	14-AUG-2000;	2000US-0225758.	PR	08-NOV-2000;	2000US-0246610.
PR	14-AUG-2000;	2000US-0225759.	PR	08-NOV-2000;	2000US-0246611.
PR	18-AUG-2000;	2000US-0226279.	PR	08-NOV-2000;	2000US-0246613.
PR	22-AUG-2000;	2000US-0226681.	PR	17-NOV-2000;	2000US-0249207.
PR	22-AUG-2000;	2000US-0226868.	PR	17-NOV-2000;	2000US-0249208.
PR	23-AUG-2000;	2000US-0227182.	PR	17-NOV-2000;	2000US-0249209.
PR	30-AUG-2000;	2000US-0228924.	PR	17-NOV-2000;	2000US-0249211.
PR	01-SEP-2000;	2000US-0229287.	PR	17-NOV-2000;	2000US-0249212.
PR	01-SEP-2000;	2000US-0229343.	PR	17-NOV-2000;	2000US-0249213.
PR	01-SEP-2000;	2000US-0229344.	PR	17-NOV-2000;	2000US-0249214.
PR	01-SEP-2000;	2000US-0229345.	PR	17-NOV-2000;	2000US-0249215.
PR	05-SEP-2000;	2000US-0229509.	PR	17-NOV-2000;	2000US-0249216.
PR	05-SEP-2000;	2000US-0229513.	PR	17-NOV-2000;	2000US-0249217.
PR	06-SEP-2000;	2000US-0230437.	PR	17-NOV-2000;	2000US-0249218.
PR	06-SEP-2000;	2000US-0230438.	PR	17-NOV-2000;	2000US-0249244.
PR	08-SEP-2000;	2000US-0231242.	PR	17-NOV-2000;	2000US-0249245.
PR	08-SEP-2000;	2000US-0231243.	PR	17-NOV-2000;	2000US-0249264.
PR	08-SEP-2000;	2000US-0231244.	PR	17-NOV-2000;	2000US-0249265.
PR	08-SEP-2000;	2000US-0231413.	PR	17-NOV-2000;	2000US-0249297.
PR	08-SEP-2000;	2000US-0231414.	PR	17-NOV-2000;	2000US-0249299.
PR	08-SEP-2000;	2000US-0232080.	PR	17-NOV-2000;	2000US-0250160.
PR	08-SEP-2000;	2000US-0232081.	PR	01-DEC-2000;	2000US-0250391.
PR	12-SEP-2000;	2000US-0231968.	PR	01-DEC-2000;	2000US-0250391.
PR	14-SEP-2000;	2000US-0232397.	PR	05-DEC-2000;	2000US-0251030.
PR	14-SEP-2000;	2000US-0232398.	PR	05-DEC-2000;	2000US-0251030.
PR	14-SEP-2000;	2000US-0232399.	PR	06-DEC-2000;	2000US-0251479.
PR	14-SEP-2000;	2000US-0232400.	PR	08-DEC-2000;	2000US-0251856.
PR	14-SEP-2000;	2000US-0232401.	PR	08-DEC-2000;	2000US-0251868.
PR	14-SEP-2000;	2000US-0233063.	PR	08-DEC-2000;	2000US-0251869.
PR	14-SEP-2000;	2000US-0233064.	PR	08-DEC-2000;	2000US-0251989.
PR	14-SEP-2000;	2000US-0233065.	PR	08-DEC-2000;	2000US-0251990.
PR	21-SEP-2000;	2000US-0234223.	PR	11-DEC-2000;	2000US-0254097.
PR	21-SEP-2000;	2000US-0234274.	PR	05-JAN-2001;	2001US-0259678.
PR	25-SEP-2000;	2000US-0234997.	XX		
PR	25-SEP-2000;	2000US-0234998.	XX		
PR	26-SEP-2000;	2000US-0235484.	XX		
PR	26-SEP-2000;	2000US-0235484.	XX		
PR	27-SEP-2000;	2000US-0235834.	XX		
PR	27-SEP-2000;	2000US-0235836.	XX		
PR	29-SEP-2000;	2000US-0236327.	XX		
PR	29-SEP-2000;	2000US-0236327.	XX		
PR	29-SEP-2000;	2000US-0236368.	XX		
PR	29-SEP-2000;	2000US-0236369.	XX		
PR	29-SEP-2000;	2000US-0236370.	XX		
PR	29-SEP-2000;	2000US-0236802.	XX		
PR	02-OCT-2000;	2000US-0237037.	XX		
PR	02-OCT-2000;	2000US-0237038.	XX		
PR	02-OCT-2000;	2000US-0237039.	XX		
PR	02-OCT-2000;	2000US-0237040.	XX		
PR	13-OCT-2000;	2000US-0239935.	XX		
PR	13-OCT-2000;	2000US-0239937.	XX		
PR	20-OCT-2000;	2000US-0241221.	XX		
PR	20-OCT-2000;	2000US-0241221.	XX		
PR	20-OCT-2000;	2000US-0241785.	XX		
PR	20-OCT-2000;	2000US-0241786.	XX		
PR	20-OCT-2000;	2000US-0241787.	XX		

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;
WPI; 2001-465566/50.
P-PSDB; AAU23008.

Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases

Claim 4; SEQ ID NO 104; 1180pp; English.

The present invention relates to the isolation of novel human enzyme polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of a wide range of

disorders including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. influenza). The polynucleotides of the invention can also be used in gene therapy. AAS40785-AAS41684 represent cDNA sequences encoding for the novel human enzyme polypeptides of the invention.

CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO CC at ftp.wipo.int/pub/published_pct_sequences.

XX

Sequence 2492 BP; 617 A; 621 C; 693 G; 558 T; 3 other;

Query Match 99.9%; Score 2450.6; DB 22; Length 2492;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2450; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTGTGTGTCTCCCTGCGGCGCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGGAAACATCGG 61
 DB |||||

QY 2 GTGTGTGTCTCCCTGCGGCGCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGGAAACATCGG 61
 DB |||||

QY 62 GCACCATGAGCGGTGCGGGCTCTTCTGCGCACACAGGCTGCGGCTGCGCTGCGGG 121
 DB |||||

QY 62 GCACCATGAGCGGTGCGGGCTCTTCTGCGCACACAGGCTGCGGCTGCGCTGCGGG 121
 DB |||||

QY 122 GTCTGTGTCTCTACCGCAACGCGGCTACTGCGCACACAGGCTGCGGCTGCGGCTGCGGCT 181
 DB |||||

QY 122 GTCTGTGTCTCTACCGCAACGCGGCTACTGCGCACACAGGCTGCGGCTGCGGCTGCGGCT 181
 DB |||||

QY 182 TCGCCTAAGAGCTTTTCTAGGCAAAATCAAGAGGAAAGAGTTTCCCATTTCCAGAAG 241
 DB |||||

QY 182 TCGCCTAAGAGCTTTTCTAGGCAAAATCAAGAGGAAAGAGTTTCCCATTTCCAGAAG 241
 DB |||||

QY 242 TTAGCCCAAGATGAATTAATGAATCAATCAGTTCTTGGACCCCGTGGAAATCTTCA 301
 DB |||||

QY 242 TTAGCCCAAGATGAATTAATGAATCAATCAGTTCTTGGACCCCGTGGAAATCTTCA 301
 DB |||||

QY 302 CTGAAGAGTGGACTCCCGAAATTTGACAGGAGGAGGAAATCCAGATGAATCTTGG 361
 DB |||||

QY 302 CTGAAGAGTGGACTCCCGAAATTTGACAGGAGGAGGAAATCCAGATGAATCTTGG 361
 DB |||||

QY 362 AGAAATGAAGAGCTAGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCTGG 421
 DB |||||

QY 362 AGAAATGAAGAGCTAGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCTGG 421
 DB |||||

QY 422 GCTTCTCCCAACACCATGTACTCAGACTAGGAGGAGATCATCAGATGGATGGTCCATCA 481
 DB |||||

QY 422 GCTTCTCCCAACACCATGTACTCAGACTAGGAGGAGATCATCAGATGGATGGTCCATCA 481
 DB |||||

QY 482 CTGTGACCTGGCAGCGCACAGGCTATTTGGGCTCAAGGGATCATCTTGGCTGGCACTG 541
 DB |||||

QY 482 CTGTGACCTGGCAGCGCACAGGCTATTTGGGCTCAAGGGATCATCTTGGCTGGCACTG 541
 DB |||||

QY 542 AGGAGCAGAAAGCCAAATCTTGCCTAACTGGCGTCCGGGAGCAGACACTAA 601
 DB |||||

QY 542 AGGAGCAGAAAGCCAAATCTTGCCTAACTGGCGTCCGGGAGCAGACACTAA 601
 DB |||||

QY 602 GCCTCAGCGAGCCAGCTAGTGGAGCGATGAGCGCTCAATCCGAGCAGAGCCACACTAA 661
 DB |||||

QY 602 GCCTCAGCGAGCCAGCTAGTGGAGCGATGAGCGCTCAATCCGAGCAGAGCCACACTAA 661
 DB |||||

QY 662 GTGAGCAAGAGAGCTACATCTCAATGCTCCAGGCTGATTAATGGAGGAC 721
 DB |||||

QY 662 GTGAGCAAGAGAGCTACATCTCAATGCTCCAGGCTGATTAATGGAGGAC 721
 DB |||||

QY 722 TGGCCAAATATTTTACTGTGTTCAGAGAGCTGAGGCTGTTGATTTCTGATGGATCAGTA 781
 DB |||||

QY 722 TGGCCAAATATTTTACTGTGTTCAGAGAGCTGAGGCTGTTGATTTCTGATGGATCAGTA 781
 DB |||||

QY 782 AAGACAAATACAGCATTCTAGTAGAAGAGACTTTGTTGGAGTCACTAATGGGAAAC 841
 DB |||||

DB |||||

QY 782 AAGACAAATACAGCATTCTAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAAAC 841
 DB |||||

QY 842 CCCAAGATAAATTAGGCATTTGGGGCTCCCAACACTTTGTGAAGTCCATTTTGAAGAACCA 901
 DB |||||

QY 842 CCCAAGATAAATTAGGCATTTGGGGCTCCCAACACTTTGTGAAGTCCATTTTGAAGAACCA 901
 DB |||||

QY 902 AGATACCTGTGGAAAAACATCTTTGAGAGAGTCCGAGAGTGGTAAAGTGGGCCATCAACA 961
 DB |||||

QY 902 AGATACCTGTGGAAAAACATCTTTGAGAGAGTCCGAGAGTGGTAAAGTGGGCCATCAACA 961
 DB |||||

QY 962 TCTCTAACAGCGCGCGGCTTTCAGATGGGCGAGCTGCTGGCTGGCTCAAGAGATTGA 1021
 DB |||||

QY 962 TCTCTAACAGCGCGCGGCTTTCAGATGGGCGAGCTGCTGGCTGGCTCAAGAGATTGA 1021
 DB |||||

QY 1022 TTGAATGACTGCTGAGTACGCTGCACAAGGAAACAGTTTAAACAGAGGCTCAGTGAAT 1081
 DB |||||

QY 1022 TTGAATGACTGCTGAGTACGCTGCACAAGGAAACAGTTTAAACAGAGGCTCAGTGAAT 1081
 DB |||||

QY 1082 TTGGATTGATTTCAGGAGAAATTTGCACCTGATGGCTCACAAGGCTTACGCTCATGGAGATA 1141
 DB |||||

QY 1082 TTGGATTGATTTCAGGAGAAATTTGCACCTGATGGCTCACAAGGCTTACGCTCATGGAGATA 1141
 DB |||||

QY 1142 TGACCTTACTCACACAGCGGATGCTGGACCAACCTGGCTTTCCGAGCTGCTCCATCGAGG 1201
 DB |||||

QY 1142 TGACCTTACTCACACAGCGGATGCTGGACCAACCTGGCTTTCCGAGCTGCTCCATCGAGG 1201
 DB |||||

QY 1202 CAGCCATGTTGAAGTGTTCAGCTCCGAGCGGCTGGCAGTGTGTGAGTGGAGCGCTGC 1261
 DB |||||

QY 1202 CAGCCATGTTGAAGTGTTCAGCTCCGAGCGGCTGGCAGTGTGTGAGTGGAGCGCTGC 1261
 DB |||||

QY 1262 AGATCCTCGGGGCTTGGGCTACACAGGAGACTATCCGTACGAGCGCATACTGCTGACA 1321
 DB |||||

QY 1262 AGATCCTCGGGGCTTGGGCTACACAGGAGACTATCCGTACGAGCGCATACTGCTGACA 1321
 DB |||||

QY 1322 CCGCATCCTCCTCATCTTTCGAGGGAACCAATGAGATTCTCCGAGTACATGCCCTGA 1381
 DB |||||

QY 1322 CCGCATCCTCCTCATCTTTCGAGGGAACCAATGAGATTCTCCGAGTACATGCCCTGA 1381
 DB |||||

QY 1382 CCGGTCTGCAGCATGCGCGGCGCATCTGACTACAGGATCCATGAGCTTAAACAGGCCA 1441
 DB |||||

QY 1382 CCGGTCTGCAGCATGCGCGGCGCATCTGACTACAGGATCCATGAGCTTAAACAGGCCA 1441
 DB |||||

QY 1442 AAGTGAGCAGCATGATGATACGCTTGGCGGAGGCTTCCGGACTCCCTGGGCGGAGCTG 1501
 DB |||||

QY 1442 AAGTGAGCAGCATGATGATACGCTTGGCGGAGGCTTCCGGACTCCCTGGGCGGAGCTG 1501
 DB |||||

QY 1502 TGGACCTGGGCTGACAGGCAACCATGAGTTGTGACCCCGAGTCTTTCGAGCAGTGGCCA 1561
 DB |||||

QY 1502 TGGACCTGGGCTGACAGGCAACCATGAGTTGTGACCCCGAGTCTTTCGAGCAGTGGCCA 1561
 DB |||||

QY 1562 ACAAGTTTGAAGGAGAACACCTACTGCTTGGCGGAGGCTTCCGAGCAGTGGCTCCGCT 1621
 DB |||||

QY 1562 ACAAGTTTGAAGGAGAACACCTACTGCTTGGCGGAGGCTTCCGAGCAGTGGCTCCGCT 1621
 DB |||||

QY 1622 TTGGCAAGACCATCATGGAGGAGCAGTGGTACTGAAGCGGGTGGGCAACATCTCATCA 1681
 DB |||||

QY 1622 TTGGCAAGACCATCATGGAGGAGCAGTGGTACTGAAGCGGGTGGGCAACATCTCATCA 1681
 DB |||||

QY 1682 ACCTGTATGGCATGAGCGGCTGCTGCGGGGCGGAGCGGCTCCATCCGATTTGGGCTCC 1741
 DB |||||

QY 1682 ACCTGTATGGCATGAGCGGCTGCTGCGGGGCGGAGCGGCTCCATCCGATTTGGGCTCC 1741
 DB |||||

QY 1742 GCAACCCAGCAGGAGTCTTCTTGGCCAAACACTTCTCGCTGGAGGCTTACTTTCAGTA 1801
 DB |||||

QY 1742 GCAACCCAGCAGGAGTCTTCTTGGCCAAACACTTCTCGCTGGAGGCTTACTTTCAGTA 1801
 DB |||||

QY 1802 ATCTTTCAGGCTTCTCAGCTGGAGCAAGTATGCTCCAGAAACCTTAGATGAGCAGATTA 1861
 DB |||||

QY 1802 ATCTTTCAGGCTTCTCAGCTGGAGCAAGTATGCTCCAGAAACCTTAGATGAGCAGATTA 1861
 DB |||||

QY 1862 AGAAAGTGTCCAGCAGATCTTTCAGAGCGGAGCTTATCTGTGCCCAACCTCTTGGACA 1921
 DB |||||

Db	181	AAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGTTTCCCATTTCCAGAAGTTAGCC	240	QY	1328	TCCTCTCTATCTTCGAGGGAACCAATCAGATTCTCCGGATGTACATCGCCCTGACGGGTC	1387
QY	248	AAAGTGAACCTTAATGAATCAATCAGTTCTTGGGACCCGTGGAAAAATTTCTTCACAG	307	Db	1321	TCCTCTCTATCTTCGAGGGAACCAATCAGATTCTCCGGATGTACATCGCCCTGACGGGTC	1380
Db	241	AGATGAACTTAATGAATCAATCAGTTCTTGGGACCCGTGGAAAAATTTCTTCACAG	300	QY	1388	TGACGATCGCGGCCCATCTCTGACTACCAAGGATCCATGAGCTTAAACAGGCAAAAGTGA	1447
QY	308	AGGTGGACTCCGAAAAATTGACCAAGGAGGAAAAATCCAGAGATGAACATTTTGAGAAAT	367	Db	1381	TGACGATCGCGGCCCATCTCTGACTACCAAGGATCCATGAGCTTAAACAGGCAAAAGTGA	1440
Db	301	AGGTGGACTCCGAAAAATTGACCAAGGAGGAAAAATCCAGAGATGAACATTTTGAGAAAT	360	QY	1448	GCACAGTCAATGGATACCGTTGGCCGAGGCTTCGGGACTCCCTGGGCGGAATGTGGACC	1507
QY	368	TGAACAGCCTAGGGCTTTTGGGCTGCAAGTCCAGAGAATAATGGTGGCTGGGCTTCT	427	Db	1441	GCACAGTCAATGGATACCGTTGGCCGAGGCTTCGGGACTCCCTGGGCGGAATGTGGACC	1500
Db	361	TGAACAGCCTAGGGCTTTTGGGCTGCAAGTCCAGAGAATAATGGTGGCTGGGCTTCT	420	QY	1508	TGGGGTGCAGGCAACCATGGAGTTGTACCCAGCTTCTCGGACAGTGCACAACAGT	1567
QY	428	CCACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCATCACTGTGA	487	Db	1501	TGGGGTGCAGGCAACCATGGAGTTGTACCCAGCTTCTCGGACAGTGCACAACAGT	1560
Db	421	CCACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCATCACTGTGA	480	QY	1568	TTGAGGAGAACACCTACTGCTTCGGCCGGACCGTGGAGACACTGCTGCTCGCTTTGGCA	1627
QY	488	CCCTGGCAGCCAGGCTATTTGGCTCAAGGGGATCATCTTGGCTGGCACTGAGGAGC	547	Db	1561	TTGAGGAGAACACCTACTGCTTCGGCCGGACCGTGGAGACACTGCTGCTCGCTTTGGCA	1620
Db	481	CCCTGGCAGCCAGGCTATTTGGCTCAAGGGGATCATCTTGGCTGGCACTGAGGAGC	540	QY	1628	AGACCATCATGGAGGAGCAGCTGGTACTGAAGCGGGTGGCCAAACATCTCATCAACCTGT	1687
QY	548	AGAAAGCCAAATACTTGCTTAACTGGCTCCGGGAGGACATTTGCAGCCTTCTGCTCA	607	Db	1621	AGACCATCATGGAGGAGCAGCTGGTACTGAAGCGGGTGGCCAAACATCTCATCAACCTGT	1680
Db	541	AGAAAGCCAAATACTTGCTTAACTGGCTCCGGGAGGACATTTGCAGCCTTCTGCTCA	600	QY	1688	ATGGCATGACGGCCCTGCTGTCGGCCGGCCAGCGCTCCATCCGATTTGGGCTCGGCAACC	1747
QY	608	CGGAGCCAGCCAGTGGGAGGATGACGCCCTCAATCCGGAGGACAGCCACACTAAGTGAAG	667	Db	1681	ATGGCATGACGGCCCTGCTGTCGGCCGGCCAGCGCTCCATCCGATTTGGGCTCGGCAACC	1740
Db	601	CGGAGCCAGCCAGTGGGAGGATGACGCCCTCAATCCGGAGGACAGCCACACTAAGTGAAG	660	QY	1748	ACGACACAGAGTTCTCTTTGGCCAAACACTTCTCGGTGGAAAGCTTACTTTGAGAAATCTCT	1800
QY	668	ACAAGAAGCACTACATCCTCAATGGCTCCAGGCTCAAGGCTGGATTAATTTGAGGAGTGGCCA	727	Db	1741	ACGACACAGAGTTCTCTTTGGCCAAACACTTCTCGGTGGAAAGCTTACTTTGAGAAATCTCT	1800
Db	661	ACAAGAAGCACTACATCCTCAATGGCTCCAGGCTCAAGGCTGGATTAATTTGAGGAGTGGCCA	720	QY	1808	TCAGCCTCTCTCAGCTGGACAAAGTATGCTCCAGAAAACCTAGATGAGCAGATTAAAGAAAG	1867
QY	728	ATAATTTTACTGTGTTGCAAGACTGAGGTGCTTGATTTCTGATGATCAGTGAAGACA	787	Db	1801	TCAGCCTCTCTCAGCTGGACAAAGTATGCTCCAGAAAACCTAGATGAGCAGATTAAAGAAAG	1860
Db	721	ATAATTTTACTGTGTTGCAAGACTGAGGTGCTTGATTTCTGATGATCAGTGAAGACA	780	QY	1868	TGTCCTCCACAGATCTTTGAGAAAGGAGCCTATATCTGTGCCACCCCTCTGACAGGACAT	1927
QY	788	AAATCAGACATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATTTGGAACCCGGAAG	847	Db	1861	TGTCCTCCACAGATCTTTGAGAAAGGAGCCTATATCTGTGCCACCCCTCTGACAGGACAT	1920
Db	781	AAATCAGACATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATTTGGAACCCGGAAG	840	QY	1928	GCTGAGGAGGAGGACAGTGTCCCTGCTACCGCCCGCCCTTACCCTAGGCGGCGGCTGCTGG	1987
QY	848	ATAAATAGGCAATTCGGGGCTCCACACACTTGTCAAGTCCATTTTGAACACCAAGATAC	907	Db	1921	GCTGAGGAGGAGGACAGTGTCCCTGCTACCGCCCGCCCTTACCCTAGGCGGCGGCTGCTGG	1980
Db	841	ATAAATAGGCAATTCGGGGCTCCACACACTTGTCAAGTCCATTTTGAACACCAAGATAC	900	QY	1988	ATGACTGTTACTCTTTTTCAGAAAGGTTGGGATTAACAGAGTTAAGCCTTTTGTTC	2047
QY	908	CTGTGGAACACATCTTTGGAGAGTCCGAGATGGGTTAAGGTGGCCATGAACATCCTCA	967	Db	1981	ATGACTGTTACTCTTTTTCAGAAAGTGTGGGATTAACAGAGTTAAGCCTTTTGTTC	2040
Db	901	CTGTGGAACACATCTTTGGAGAGTCCGAGATGGGTTAAGGTGGCCATGAACATCCTCA	960	QY	2048	CCGTCTGCACCTGAAGGGTTGTCGCCCTGGCTGGGAGGAGCCTTCCAGGTTTGTGACCTG	2107
QY	968	ACAGCGCGGTTTACAGCATGGGAGGCTCGTGGCTGGGCTGCTCAAGAGATTGATGAAA	1027	Db	2041	CCGTCTGCACCTGAAGGGTTGTCGCCCTGGCTGGGAGGAGCCTTCCAGGTTTGTGACCTG	2100
Db	961	ACAGCGCGGTTTACAGCATGGGAGGCTCGTGGCTGGGCTGCTCAAGAGATTGATGAAA	1020	QY	2108	CAGCAGTGTCTCTAACAGGACCATCACAGCTTCTGAAGTGGAGCGGAGAGAGAGATG	2167
QY	1028	TGACTGCTGAGTACGCTGCACAAGGAAACAGTTTAAAGAGGCTCAGTGAATTTGGAT	1087	Db	2101	CAGCAGTGTCTCTAACAGGACCATCACAGCTTCTGAAGTGGAGCGGAGAGAGAGATG	2160
Db	1021	TGACTGCTGAGTACGCTGCACAAGGAAACAGTTTAAAGAGGCTCAGTGAATTTGGAT	1080	QY	2168	GAATTTGCTGACCCCTGAACTGGCGGGTATTTCTGGTCAATTGAGGAGACACCATAGTGGAA	2227
QY	1088	TGATTCAGAGAAATTTGACATGATGGCTCAGAAGGCTTACGTCATGGAGAGTATGACCT	1147	Db	2161	GAATTTGCTGACCCCTGAACTGGCGGGTATTTCTGGTCAATTGAGGAGACACCATAGTGGAA	2220
Db	1081	TGATTCAGAGAAATTTGACATGATGGCTCAGAAGGCTTACGTCATGGAGAGTATGACCT	1140	QY	2228	ACTGGGGCTTATGCTGCTGCTCCAGGCTGAGGTGGGTGGGAGCCTGCTGTCAGGTTG	2287
QY	1148	ACCTCACACAGGATGCTGGACCAACCTGGCTTTCCGACTGCTCCATCGAGGAGGACCA	1207	Db	2221	ACTGGGGCTTATGCTGCTGCTCCAGGCTGAGGTGGGTGGGAGCCTGCTGTCAGGTTG	2280
Db	1141	ACCTCACACAGGATGCTGGACCAACCTGGCTTTCCGACTGCTCCATCGAGGAGGACCA	1200	QY	2288	GATAGCCATTTGCTCAACACACATTTCTTAAAGAAACAGCTTGAAGCCTCTGTCTGGG	2347
QY	1208	TGGTGAAGTGTTCAGCTCCGAGGCGGCTGGCAGTGTGAGTGAGGCGCTGAGATCC	1267	Db	2281	GATAGCCATTTGCTCAACACACATTTCTTAAAGAAACAGCTTGAAGCCTCTGTCTGGG	2340
Db	1201	TGGTGAAGTGTTCAGCTCCGAGGCGGCTGGCAGTGTGAGTGAGGCGCTGAGATCC	1260	QY	2348	TCATTTCAATTTAAACTAGAAGCAGGACACTTAAACATGTACAGGAAACCATTTTAAACAA	2407
QY	1268	TCGGGGGCTTGGGCTACACAGAGGAGTATCCGTACGAGCGCATACTGGGTGACACCCGCA	1327	Db	2341	TCATTTCAATTTAAACTAGAAGCAGGACACTTAAACATGTACAGGAAACCATTTTAAACAA	2400
Db	1261	TCGGGGGCTTGGGCTACACAGAGGAGTATCCGTACGAGCGCATACTGGGTGACACCCGCA	1320				

QY	2408	GAATATAAATGTCAAACTCTGTACTCTTAAATAAAAAA	2447
Db	2401	GAATATAAATGTCAAACTCTGTACTCTTAAATAAAAAA	2440
RESULT 8			
ID	AAH16781	AAH16781 standard; cDNA; 2399 BP.	
XX	AAH16781;		
XX	AC		
DT	26-JUN-2001	(first entry)	
XX	Human cDNA sequence	SEQ ID NO:16009.	
DE	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.		
KW	Homo sapiens.		
OS	EPI074617-A2.		
PN	07-FEB-2001.		
PD	28-JUL-2000; 2000EP-0116126.		
XX	29-JUL-1999; 99JP-0248036.		
PR	27-AUG-1999; 99JP-0300253.		
PR	11-JAN-2000; 2000JP-0118776.		
PR	02-MAY-2000; 2000JP-0183767.		
PR	09-JUN-2000; 2000JP-0241899.		
XX	(HELI-) HELIX RES INST.		
PA	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;		
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;		
PI	WPI; 2001-318749/34.		
DR	Primer sets for synthesizing polynucleotides, particularly the 5602		
XX	full-length cDNAs defined in the specification, and for the detection		
PT	and/or diagnosis of the abnormality of the proteins encoded by the		
PT	full-length cDNAs		
XX	Claim 8; SEQ ID 16009; 2537pp + CD ROM; English.		
PS	The present invention describes primer sets for synthesising 5602		
XX	full-length cDNAs defined in the specification. Where a primer set		
CC	comprises: (a) an oligo-dT primer and an oligonucleotide complementary		
CC	to the complementary strand of a polynucleotide which comprises one of		
CC	the 5602 nucleotide sequences defined in the specification, where the		
CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination		
CC	of an oligonucleotide comprising a sequence complementary to the		
CC	complementary strand of a polynucleotide which comprises a 5'-end		
CC	sequence and an oligonucleotide comprising a sequence complementary to a		
CC	polynucleotide which comprises a 3'-end sequence, where the		
CC	oligonucleotide comprises at least 15 nucleotides and the combination of		
CC	the 5'-end sequence/3'-end sequence is selected from those defined in		
CC	the specification. The primer sets can be used in antisense therapy and		
CC	in gene therapy. The primers are useful for synthesising polynucleotides,		
CC	particularly full-length cDNAs. The primers are also useful for the		
CC	detection and/or diagnosis of the abnormality of the proteins encoded by		
CC	the full-length cDNAs. The primers allow obtaining of the full-length		
CC	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and		
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to		
CC	AAH93893 represent human amino acid sequences; and AAH13629 to AAH13632		
CC	represent oligonucleotides, all of which are used in the exemplification		
CC	of the present invention.		
XX	Sequence 2399 BP; 597 A; 606 C; 672 G; 524 T; 0 other;		
XX	Query Match 97.2%; Score 2382.2; DB 22; Length 2399;		
XX	Best Local Similarity 99.8%; Pred. No. 0;		
XX	Matches 2396; Conservative 0; Mismatches 3; Indels 2; Gaps 1;		

Qy	21	GCTAAGAGGGGAGACTGAGCTGAGGCTGGGGAACATCGGGCAGCATGAGCGCTCGG	80
Db	1	GCTAAGAGGGGAGACTGAGCTGAGGCTGGGGAACATCGGGCAGCATGAGCGCTCGG	60
Qy	81	GCTCTTCTGCGCACCGCTGCGGCTGCTGCTGCGGGGCTGCTGCTGCTTACCGC	140
Db	61	GCTCTTCTGCGCACCGCTGCGGCTGCTGCTGCGGGGCTGCTGCTGCTTACCGC	120
Qy	141	GAACCGCGGCTTACTGCGCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTTTCCT	200
Db	121	GAACCGCGGCTTACTGCGCACCGCTGCTGCTGCTGCTGCTGCTGCTTTCCT	180
Qy	201	AGGCAAAATCAAGAAGAAGAAGTTTCCCATTTCCAGAGCTAGCCAAAGATGAAT	260
Db	181	AGGCAAAATCAAGAAGAAGAAGTTTCCCATTTCCAGAGCTAGCCAAAGATGAAT	240
Qy	261	TGAATATCAATCAGTTCTTGGGACCGCTGGAATAATTTCTTCTGAGAGGCTG	320
Db	241	TGAATATCAATCAGTTCTTGGGACCGCTGGAATAATTTCTTCTGAGAGGCTG	300
Qy	321	AAAAATTGACAGGAGGAAATCCAGATGAATTTGGAGAAATTTGAAGAGCTAGG	380
Db	301	AAAAATTGACAGGAGGAAATCCAGATGAATTTGGAGAAATTTGAAGAGCTAGG	360
Qy	381	GCTTTTGGGCTGCAAGTCCCAAGAGATATGCTGGCTGCTTCTTCTCAACACCAT	440
Db	361	GCTTTTGGGCTGCAAGTCCCAAGAGATATGCTGGCTGCTTCTTCTCAACACCAT	420
Qy	441	CTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCATCTGTACCTGGCAG	500
Db	421	CTCAGACTAGGGAGATCATCAGCATGGATGGTCCATCATCTGTACCTGGCAG	480
Qy	501	CCAGGCTATGGCTCAAGGGATCATCTTGGCTGCTGCTGCTGCTGCTGCTGCTG	560
Db	481	CCAGGCTATGGCTCAAGGGATCATCTTGGCTGCTGCTGCTGCTGCTGCTGCTG	540
Qy	561	CTTGCCTAAACTGGCTCCGGGAGCACATTCAGCCTTCTGCTTCTGCTGCTGCTG	620
Db	541	CTTGCCTAAACTGGCTCCGGGAGCACATTCAGCCTTCTGCTTCTGCTGCTGCTG	600
Qy	621	TGGAGGCTGAGGCTCAATCGGAGCAGCCACACTAAGTGAAGCAAGACACAT	680
Db	601	TGGAGGCTGAGGCTCAATCGGAGCAGCCACACTAAGTGAAGCAAGACACAT	660
Qy	681	CATCTCTCAATGGCTCCAAAGTCTGATTAATGAGGAGCTGGCCAAATATTTT	740
Db	661	CATCTCTCAATGGCTCCAAAGTCTGATTAATGAGGAGCTGGCCAAATATTTT	720
Qy	741	GTTTGCAAGACTGAGCTGCTGATTTCTGATGATCAGTCAAGACAAATACACAT	800
Db	721	GTTTGCAAGACTGAGCTGCTGATTTCTGATGATCAGTCAAGACAAATACACAT	780
Qy	801	CATAGTGAAGAGACTTTGCTGGAGTCACTAATGGAACCCGAGATAAATAGG	860
Db	781	CATAGTGAAGAGACTTTGCTGGAGTCACTAATGGAACCCGAGATAAATAGG	840
Qy	861	TCGGGCTCCAAACTTTGTGAAGTCCATTTTGAACACCAAGATACCTTGTGA	920
Db	841	TCGGGCTCCAAACTTTGTGAAGTCCATTTTGAACACCAAGATACCTTGTGA	900
Qy	921	CTTTGAGAGGCTCGAGATGGTTTAAAGTGGCCATCAACATCTTCAACAGCG	980
Db	901	CTTTGAGAGGCTCGAGATGGTTTAAAGTGGCCATCAACATCTTCAACAGCG	960
Qy	981	CAGCATGGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1040
Db	961	CAGCATGGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1020
Qy	1041	CGCTGTCACAGGAAACAGTTTACAGAGGCTCAGTGAATTTGATTTGATTT	1100
Db	1021	CGCTGTCACAGGAAACAGTTTACAGAGGCTCAGTGAATTTGATTTGATTT	1080

```
QY 1101 ATTTGACTATGCTCAGAGGCTTACGTATGAGAGTATGACCTACCTACAGCAGG 1160
DB 1081 ATTTGACTATGCTCAGAGGCTTACGTATGAGAGTATGACCTACCTACAGCAGG 1140
QY 1161 GATGCTGGACCAACCTGGCTTCCGACTGCTCCATCGAGGAGCCATGTTGAAGTGT 1220
DB 1141 GATGCTGGACCAACCTGGCTTCCGACTGCTCCATCGAGGAGCCATGTTGAAGTGT 1200
QY 1221 CAGCTCCGAGCGCCCTGGCAGTGTGAGTGGAGCGCTGCAGATCCTCGGGGCTTGGG 1280
DB 1201 CAGCTCCGAGCGCCCTGGCAGTGTGAGTGGAGCGCTGCAGATCCTCGGGGCTTGGG 1260
QY 1281 CTACACAGGAGACTATCCGTACGAGGCGATPACTGCGTGACACCCGATCCTCCATCTT 1340
DB 1261 CTACACAGGAGACTATCCGTACGAGGCGATPACTGCGTGACACCCGATCCTCCATCTT 1320
QY 1341 CGAGGGAACCAATGAGATCTCCGATGTACATCGCCCTGACGGGTCTGAGCATGCCGG 1400
DB 1321 CGAGGGAACCAATGAGATCTCCGATGTACATCGCCCTGACGGGTCTGAGCATGCCGG 1380
QY 1401 CCGCATCTGACTACGAGGATCCATGAGCTTAAACAGGCCAAAGTGAGCACAGTCAATGA 1460
DB 1381 CCGCATCTGACTACGAGGATCCATGAGCTTAAACAGGCCAAAGTGAGCACAGTCAATGA 1440
QY 1461 TACCGTTGGCCGAGGCTTCGGGACTCCCTGGGCCGAACTGTGACCTGGGGCTGACAGG 1520
DB 1441 TACCGTTGGCCGAGGCTTCGGGACTCCCTGGGCCGAACTGTGACCTGGGGCTGACAGG 1500
QY 1521 CAACCATGGAGTTGTCACCCAGTCTCGGGACAGTCCCAACAAAGTTTGAGGAGAACAC 1580
DB 1501 CAACCATGGAGTTGTCACCCAGTCTCGGGACAGTCCCAACAAAGTTTGAGGAGAACAC 1560
QY 1581 CTACTGCTTGGCCGAGGCTTGAGACACTGCTGCTCCGCTTTGGCAAGCAACATCATGA 1640
DB 1561 CTACTGCTTGGCCGAGGCTTGAGACACTGCTGCTCCGCTTTGGCAAGCAACATCATGA 1620
QY 1641 GGAGCAGCTGGTACTCAACGGGGTGGCCAAACATCTCTATCAACCTGTATGGCATACGGC 1700
DB 1621 GGAGCAGCTGGTACTCAACGGGGTGGCCAAACATCTCTATCAACCTGTATGGCATACGGC 1680
QY 1701 CGTGCTGTCGCGGGCAGCGCTCCATCGGATGGGCTCCGCAACCAAGTGTATGGCATACGGC 1760
DB 1681 CGTGCTGTCGCGGGCAGCGCTCCATCGGATGGGCTCCGCAACCAAGTGTATGGCATACGGC 1740
QY 1761 TCTCTTGGCCACACCTTCTGGTGGAAGCTTACTTGCAGAACTCTCTCAGCCTCTCTCA 1820
DB 1741 TCTCTTGGCCACACCTTCTGGTGGAAGCTTACTTGCAGAACTCTCTCAGCCTCTCTCA 1800
QY 1821 GCTGACAAAGTATGCTCCAGAAACCTTAGATGAGCAGATTAGAAGTGTCCAGCAGAT 1880
DB 1801 GCTGACAAAGTATGCTCCAGAAACCTTAGATGAGCAGATTAGAAGTGTCCAGCAGAT 1860
QY 1881 CTTTGAGAGCGAGCCTATATCTGTGCCACCCCTCTGGACAGGACATGCTGAGGAGGGG 1940
DB 1861 CTTTGAGAGCGAGCCTATATCTGTGCCACCCCTCTGGACAGGACATGCTGAGGAGGGG 1920
QY 1941 ACAGTGTCCCTGCTACCGCCCGCCCTACCCATGGCCGTTGCTGGATGACTGTACTC 2000
DB 1921 ACAGTGTCCCTGCTACCGCCCGCCCTACCCATGGCCGTTGCTGGATGACTGTACTC 1980
QY 2001 TTTTTCAGAAAGGTGTTGGGATATACAGAGTAAAGCCTTTTGTTCGCCGCTGCACCTG 2060
DB 1981 TTTTTCAGAAAGGTGTTGGGATATACAGAGTAAAGCCTTTTGTTCGCCGCTGCACCTG 2040
QY 2061 AAGGTTTCTGCGCTGGGAGAGCCTCTTCAGGTTTTCAGCTGCGAGCAGTGTCT 2120
DB 2041 AAGGTTTCTGCGCTGGGAGAGCCTCTTCAGGTTTTCAGCTGCGAGCAGTGTCT 2100
QY 2121 CTACAGGACCATCACAGCTTCTCACTGAGCCCGGAGAGAGAGATGGAATTGCTGACCC 2180
DB 2101 CTACAGGACCATCACAGCTTCTCACTGAGCCCGGAGAGAGATGGAATTGCTGACCC 2160
QY 2181 CTGGAACCTGGCGGCTATTTCTGGTCAATGAGGAGACACCAATAGTGGAACTGGGCTTATG 2240
```

```
DB 2159 CTGGAACCTGGCGGCTATTTCTGGTCAATGAGGAGACACCAATAGTGGAACTGGGCTTATG 2218
QY 2241 CTGCTCCCTCCAGGCTGTCAGGTGGGTGGGACCTGTCAGGTGTCAGGTGTCAGGTGTCAT 2300
DB 2219 CTGCTCCCTCCAGGCTGTCAGGTGGGTGGGACCTGTCAGGTGTCAGGTGTCAGGTGTCAT 2278
QY 2301 GGTCAACCAACACATCTCTTAAGAAACAGCTTGAAGAGCTGTCCTGGGTCAATTCATTTAAA 2360
DB 2279 GGTCAACCAACACATCTCTTAAGAAACAGCTTGAAGAGCTGTCCTGGGTCAATTCATTTAAA 2338
QY 2361 CTAGAGCAGAGCAGCTTAAACATGTACCAGGACCAATTTACAAAAGATATAAATGT 2420
DB 2339 CTAGAGCAGAGCAGCTTAAACATGTACCAGGACCAATTTACAAAAGATATAAATGT 2398
QY 2421 C 2421
DB 2399 C 2399

RESULT 9
AAC76009
ID AAC76009 standard; cdna; 2218 BP.
XX AAC76009;
XX AC AAC76009;
XX DT 08-FEB-2001 (first entry)
XX DE Human ORFX ORF1564 polynucleotide sequence SEQ ID NO:3127.
XX KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antilarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antineoplastic; antihypertensive;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX OS Homo sapiens.
XX PN WO200058473-A2.
XX PD 05-OCT-2000.
XX PF 31-MAR-2000; 2000WO-US08621.
XX PR 31-MAR-1999; 99US-0127607.
XX PR 02-APR-1999; 99US-0127636.
XX PR 30-APR-1999; 99US-0127728.
XX PR 30-MAR-2000; 2000US-0540763.
XX PA (CURA-) CURAGEN CORP.
XX PI Shimkets RA, Leach M;
XX DR WPI; 2000-602362/57.
XX DR P-PSDB; AAB41800.
XX PT Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
XX PS Claim 5; Page 2343-2344; 5507pp; English.
XX CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
```

sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 2218 BP; 573 A; 548 C; 609 G; 487 T; 1 other;

Query Match 90.3%; Score 2213.8; DB-21; Length 2218;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2215; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 236 CAGAAGTTAGCAAGATGAATTAATGAAATCAATCAGTTCTTTGGACCCGCGTGAATAAT 295
Db 2 CAGAAAGTTAGCAAGATGAATTAATGAAATCAATCAGTTCTTTGGACCCGCGTGAATAAT 61

Qy 296 TCTTCACTGAAGAGTGGACTCCGAAATAATGACAGGAAGGAAATCCAGATGAA 355
Db 62 TCTTCACTGAAGAGTGGACTCCGAAATAATGACAGGAAGGAAATCCAGATGAA 121

Qy 356 CTTTGGAGAATTGAAGAGCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTG 415
Db 122 CTTTGGAGAATTGAAGAGCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTG 181

Qy 416 GCCTGGGCTTCTCCAAACCATGTTACTCAAGACTAGGGAGATCATCAGATGGATGGT 475
Db 182 GCCTGGGCTTCTCCAAACCATGTTACTCAAGACTAGGGAGATCATCAGATGGATGGT 241

Qy 476 CCATCACTGTACCCCTGCGAGCGCACCGACTATTGGCTTCAAGGGGATCATCTTGCTG 535
Db 242 CCATCACTGTACCCCTGCGAGCGCACCGACTATTGGCTTCAAGGGGATCATCTTGCTG 301

Qy 536 GCATGAGGAGCAAGAAAGCCAAATPACTTGGCTTAACTGGCTCGGGGAGCAATTCAG 595
Db 302 GCATGAGGAGCAAGAAAGCCAAATPACTTGGCTTAACTGGCTCGGGGAGCAATTCAG 361

Qy 596 CCTTCTGCTCAGCGGAGCCAGTGGAGCGGATGAGCTCAATCCGGAGCAGAGCCCA 655
Db 362 CCTTCTGCTCAGCGGAGCCAGTGGAGCGGATGAGCTCAATCCGGAGCAGAGCCCA 421

Qy 656 CACTAAGTGAAGCAAGAACACTACATCTCTCAATGGCTCCAGGCTGGATTACTAATG 715
Db 422 CACTAAGTGAAGCAAGAACACTACATCTCTCAATGGCTCCAGGCTGGATTACTAATG 481

Qy 716 GAGGACTGGCCAAATATTTTACTGTGTGTTGCAAGACTGAGGTGTTGATTCATGGAT 775
Db 482 GAGGACTGGCCAAATATTTTACTGTGTGTTGCAAGACTGAGGTGTTGATTCATGGAT 541

Qy 776 CAGTCAAGACAAATACAGCATTCATAGTAGAAGAGACTTTTGGTGGAGTCACTAATG 835
Db 542 CAGTCAAGACAAATACAGCATTCATAGTAGAAGAGACTTTTGGTGGAGTCACTAATG 601

Qy 836 GGAACCCGAGATAAATTAGGCATTCGGGCTCCAACTGTGTAAGTCCATTTTGAAA 895
Db 602 GGAACCCGAGATAAATTAGGCATTCGGGCTCCAACTGTGTAAGTCCATTTTGAAA 661

Qy 896 ACACCAAGTACCTGTGGAACACATCTTGGAGAGTGGAGTGGTTAAGTGCCCA 955
Db 662 ACACCAAGTACCTGTGGAACACATCTTGGAGAGTGGAGTGGTTAAGTGCCCA 721

Qy 956 TGAACATCTCAACAGCGCGCGTTACATATGGGAGCGCTGCTGGCTGGCTGCTCAAGA 1015
Db 722 TGAACATCTCAACAGCGCGCGTTACATATGGGAGCGCTGCTGGCTGGCTGCTCAAGA 781

Qy 1016 GATTGATTGAATGACTGCTGAGTACGCTGCACAAAGAAACAGTTTAAAGAGGCTCA 1075
Db 782 GATTGATTGAATGACTGCTGAGTACGCTGCACAAAGAAACAGTTTAAAGAGGCTCA 841

Qy 1076 GTGAATTTGGATTGATTTCAGGAGAAATTTGCACTGATGGCTCAGAAGCTTTACGTCATGG 1135
Db 842 GTGAATTTGGATTGATTTCAGGAGAAATTTGCACTGATGGCTCAGAAGCTTTACGTCATGG 901

Qy 1136 AGAGTATGACCTPACCTCACAGCAGGATGTGGACCAACCTGGCTTTCCGACTGCTCCA 1195
Db 902 AGAGTATGACCTPACCTCACAGCAGGATGTGGACCAACCTGGCTTTCCGACTGCTCCA 961

Qy 1196 TCGAGGCGCCATGTGAAGGTGTTACGCTCCGAGCGCCCTGGCAGTGTGTGAGTGAAG 1255
Db 962 TCGAGGCGCCATGTGAAGGTGTTACGCTCCGAGCGCCCTGGCAGTGTGTGAGTGAAG 1021

Qy 1256 CGCTGCAGATCTCGGGGCTTTGGGCTACACAAAGGAGTATCCGTACGAGCGCATCTGCG 1315
Db 1022 CGCTGCAGATCTCGGGGCTTTGGGCTACACAAAGGAGTATCCGTACGAGCGCATCTGCG 1081

Qy 1316 GTGACACCCGATCTCTCTCATCTTCGAGGGAACCAATGAGATTTCTCGGATGTACATCG 1375
Db 1082 GTGACACCCGATCTCTCTCATCTTCGAGGGAACCAATGAGATTTCTCGGATGTACATCG 1141

Qy 1376 CCTGTACGGGCTGTGCAGCATGCGCGCGCATCTGACTACCGAGATPCCATGAGCTTAAAC 1435
Db 1142 CCTGTACGGGCTGTGCAGCATGCGCGCGCATCTGACTACCGAGATPCCATGAGCTTAAAC 1201

Qy 1436 AGGCCAAAGTGTGACAGCATGATGGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCG 1495
Db 1202 AGGCCAAAGTGTGACAGCATGATGGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCG 1261

Qy 1496 GAACTGTGACCTGGGCTGTGACAGCAACCATGAGTTGTGACCCCACTCTTCGCGACA 1555
Db 1262 GAACTGTGACCTGGGCTGTGACAGCAACCATGAGTTGTGACCCCACTCTTCGCGACA 1321

Qy 1556 GTGCCAAAGTTTGTAGGAGAACACTACTGTTTCGGCGGAGGCTTCGGGACTCCCTGGGCG 1615
Db 1322 GTGCCAAAGTTTGTAGGAGAACACTACTGTTTCGGCGGAGGCTTCGGGACTCCCTGGGCG 1381

Qy 1616 TCCGCTTTTGGCAAGACCATCATGAGGAGCAGCTGTTGTAAGCAGGCTGGCCCAACATCC 1675
Db 1382 TCCGCTTTTGGCAAGACCATCATGAGGAGCAGCTGTTGTAAGCAGGCTGGCCCAACATCC 1441

Qy 1676 TCATCAACCTGTATGTCATGACGGCGCTGTGTCGGGGGCGAGCCGCTCCATCCGCTATG 1735
Db 1442 TCATCAACCTGTATGTCATGACGGCGCTGTGTCGGGGGCGAGCCGCTCCATCCGCTATG 1501

Qy 1736 GGCTCCGACACAGCAGGAGGTTCTTCTTGGCCAAACACTTCTCGCTGGAAGCTTACT 1795
Db 1502 GGCTCCGACACAGCAGGAGGTTCTTCTTGGCCAAACACTTCTCGCTGGAAGCTTACT 1561

Qy 1796 TGCAGATCTCTTTCAGCTCTCTCAGCTGAGCAGATGCTGCAGAAAACCTAGATGAGC 1855
Db 1562 TGCAGATCTCTTTCAGCTCTCTCAGCTGAGCAGATGCTGCAGAAAACCTAGATGAGC 1621

Qy 1856 AGATTAAAGAAAGTGTCCACAGCATCTTTCAGAGGAGGCTTATATCTGTGCCACCCCTC 1915
Db 1622 AGATTAAAGAAAGTGTCCACAGCATCTTTCAGAGGAGGCTTATATCTGTGCCACCCCTC 1681

Qy 1916 TGCAGAGGACATGTCAGGCGAGGAGCAGTGTCCCTGCTTACCGCCGCGCCCTTACCCATG 1975
Db 1682 TGCAGAGGACATGTCAGGCGAGGAGCAGTGTCCCTGCTTACCGCCGCGCCCTTACCCATG 1741

Qy 1976 GCCCGTTGCTGGATGACTGTACTCTTTTTCAGAGGCTTGGGATTTATCAGAGTTAA 2035
Db 1742 GCCCGTTGCTGGATGACTGTACTCTTTTTCAGAGGCTTGGGATTTATCAGAGTTAA 1801


```
QY 2036 GCCTTTTGTCCCGTCTGCACCTGAAGGGTGTGCGCTGGCGTGGAGAGCCCTCTTCCA 2095
Db |||||||
QY 1802 GCCTTTTGTCCCGTCTGCACCTGAAGGGTGTGCGCTGGCGTGGAGAGCCCTCTTCCA 1861
Db |||||||
QY 2096 GCCTTTTGTCCCGTCTGCACCTGAAGGGTGTGCGCTGGCGTGGAGAGCCCTCTTCCA 2155
Db |||||||
QY 1862 GCCTTTTGTCCCGTCTGCACCTGAAGGGTGTGCGCTGGCGTGGAGAGCCCTCTTCCA 1921
Db |||||||
QY 2156 AGAGAGAGAAATGGAATTCCTGACCCCTGGAACCTGGCGGTATTCTGGTCAATGAGAGAC 2215
Db |||||||
QY 1922 AGAGAGAGAAATGGAATTCCTGACCCCTGGAACCTGGCGGTATTCTGGTCAATGAGAGAC 1981
Db |||||||
QY 2216 ACCATAGTGGAACTGGGGCTTATGCTGCTCCCTCCAGGGTGTGAGGTGGGTGGGACCT 2275
Db |||||||
QY 1982 ACCATAGTGGAACTGGGGCTTATGCTGCTCCCTCCAGGGTGTGAGGTGGGTGGGACCT 2041
Db |||||||
QY 2276 GTGTCAGGTGTGGATAGCAATTCCTGCTCAACCCACACATTCCTCTAAGAAACAGCTTGAAA 2335
Db |||||||
QY 2042 GTGTCAGGTGTGGATAGCAATTCCTGCTCAACCCACACATTCCTCTAAGAAACAGCTTGAAA 2101
Db |||||||
QY 2336 GCTCTGCTGGGTCAATTCATTTAACTAGAGAGAGGCACCTTAAACATGTACCAGGAA 2395
Db |||||||
QY 2102 GCTCTGCTGGGTCAATTCATTTAACTAGAGAGAGGCACCTTAAACATGTACCAGGAA 2161
Db |||||||
QY 2396 CCATTTTAAAGAAATATAAATGTCACATCTGTGTACTGTGTTAAAAAAGGAGGAA 2452
Db |||||||
QY 2162 CCATTTTAAAGAAATATAAATGTCACATCTGTGTACTGTGTTAAAAAAGGAGGAA 2218
Db |||||||
RESULT 10
AAH15804
ID AAH15804 standard; cDNA; 2400 BP.
XX
AC AAH15804;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:14270.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
XX
PR 27-AUG-1999; 99JP-0300253.
XX
PR 11-JAN-2000; 2000JP-0118776.
XX
PR 02-MAY-2000; 2000JP-0183767.
XX
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs.
XX
PS Claim 8; SEQ ID 14270; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
```

```
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
```

Seq Sequence 2400 BP; 570 A; 629 C; 685 G; 516 T; 0 other;

Query Match 86.3%; Score 2115.4; DB 22; Length 2400;

Best Local Similarity 94.2%; Pred No. 0;

Matches 2262; Conservative 0; Mismatches 56; Indels 83; Gaps 3;

```
QY 21 GCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGGAAACATCGGGCAGCATGAGCGGCTGCGG 80
Db |||||||
QY 1 GCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGGAAACATCGGGCAGCATGAGCGGCTGCGG 60
Db |||||||
QY 81 GCTCTTCTCGCGCACCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 140
Db |||||||
QY 61 GCTCTTCTCGCGCACCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 120
Db |||||||
QY 141 GAACCGCGGCTACTGCGCACCGGCTGACGAGCTGTACGAGCTTTCGCCAAAGAGCTTTTCCT 200
Db |||||||
QY 121 GAACCGCGGCTACTGCGCACCGGCTGACGAGCTGTACGAGCTTTCGCCAAAGAGCTTTTCCT 180
Db |||||||
QY 201 AGGCAAAATCAAGAGAAA----- 219
Db |||||||
QY 181 AGGCAAAATCAAGAGGTAACGGGAGCCCTGGGCGAACCTTCTGTCTGGCTCCCGCTT 240
Db -----GAAGTTTCCCATTTCCAGAAGTTACCCAGAT 252
QY 241 TTCACCTCAGTGCAGAGACTGGTGTGAATTTGTGAGATTCCTCCCAAACTGCCAGAGA 300
Db |||||||
QY 253 GAATCT-----TAATGAATCAATCAGTCTTGGGACCCGCTGGAAAAATTTCT 298
Db |||||||
QY 301 GATACACCTTGGCGCGGAGCGGTGTTAACACTCCGGATTCCTGAGTTCAGGAAACCTTT 360
Db |||||||
QY 299 TCACCTGAAGAGGTGGACTCCCGGAAATTTGACCAAGAGGAGGAAATCCAGATGAACCTT 358
Db |||||||
QY 361 CCCAGAGAAAGTGGACTCCCGGAAATTTGACCAAGAGGAGGAAATCCAGATGAACCTT 420
Db |||||||
QY 359 TGGAGAAATTTGAAGAGCTAGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCC 418
Db |||||||
QY 421 TGGAGAAATTTGAAGAGCTAGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCC 480
Db |||||||
QY 419 TGGGCTTCTCCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCA 478
Db |||||||
QY 481 TGGGCTTCTCCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCA 540
Db |||||||
QY 479 TCACCTGTACCTTGGCAGCGCCACCGGCTATTGGGCTCAAGGGGATCATCTTTGGCTGGCA 538
Db |||||||
QY 541 TCACCTGTACCTTGGCAGCGCCACCGGCTATTGGGCTCAAGGGGATCATCTTTGGCTGGCA 600
Db |||||||
QY 539 CTGAGGAGCAGAAAGCCCAATTTGCTTAACTGGCGTGGGGGAGCAGCATTCGAGCCT 598
Db |||||||
QY 601 CTGAGGAGCAGAAAGCCCAATTTGCTTAACTGGCGTGGGGGAGCAGCATTCGAGCCT 660
Db |||||||
QY 599 TCTGCTCTACGGAGCGCAGCAGTGGGAGCGATGTCAGCTCAATCCGGAGCAGAGCCACAC 658
Db |||||||
QY 661 TCTGCTCTACGGAGCGCAGCAGTGGGAGCGATGTCAGCTCAATCCGGAGCAGAGCCACAC 720
Db |||||||
```


PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0233223.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0259678.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-465566/50.
P-PSDB; AAU23012.
Novel polypeptides and polynucleotides useful for diagnosing,
preventing, treating neural, immune system, muscular, reproductive,
pulmonary, cardiovascular, renal, proliferative disorders and cancerous
diseases
Claim 4; SEQ ID NO 108; 1180pp; English.

XX The present invention relates to the isolation of novel human enzyme
CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders, including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders
CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders
CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
CC blood-related disorders (e.g. haemophilia), reproductive disorders
CC (e.g. infertility) and infectious disorders (e.g. influenza). The
CC polynucleotides of the invention can also be used in gene therapy.
CC AAS40785-AAS41684 represent cDNA sequences encoding for the novel human
CC enzyme polypeptides of the invention.
CC Note: the sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1118 BP; 245 A; 303 C; 311 G; 248 T; 11 other;

Query Match 39, 98; Score 977.2; DB 22; Length 1118;
Best Local Similarity 92.88; Pred. No. 8e-253;
Matches 1018; Conservative 2; Mismatches 77; Indels 0; Gaps 0;

QY 1015 AGATTGATGAATGACTGCTGAGTACGCTGCACAGGAACAGTTTAAACAGAGGCTC 1074
DB 4 ANATTGAKTGAATGACTGCTGAGTACGCTGCACAGGAACAGTTTAAACAGAGGCTC 63

QY 1075 AGTGAATTGATGATTCAGAGAAATTTGCACCTGATGGCTCAGAAAGGCTTACGTCATG 1134
DB 64 AGTGAATTGATGATTCAGAGAAATTTGCACCTGATGGCTCAGAAAGGCTTACGTCATG 123

QY 1135 GAGATGATGACTACCTCAGCAGGAGGATGCTGACCAACCTGGCTTTCCGGAGCTGCC 1194
DB 124 GAGATGATGACTACCTCAGCAGGAGGATGCTGACCAACCTGGCTTTCCGGAGCTGCC 183

QY 1195 ATCGAGGAGCAGTGTCAAGGTGTTACGCTCCGAGGCGCTGGCAGTGTGTGAGTGAG 1254
DB 184 ATCGAGGAGCAGTGTCAAGGTGTTACGCTCCGAGGCGCTGGCAGTGTGTGAGTGAG 243

QY 1255 GCGCTGCAGATCCTCGGGGGCTTGGGCTACACAAAGGGACTATCCGATCAGCGCATCTG 1314
DB 244 GCGCTGCAGATCCTCGGGGGCTTGGGCTACACAAAGGGACTATCCGATCAGCGCATCTG 303

QY 1315 CGTGACACCGCATCTCTCATCTTCGAGGGAACCAATGATGATTCCTCGGATGTACATC 1374
DB 304 CGTGACACCGCATCTCTCATCTTCGAGGGAACCAATGATGATTCCTCGGATGTACATC 363

QY 1375 GCGCTGACGGGTCTGCAGCATCCGCGCGCATCTCTGATACAGGATCCATGAGCTTAAA 1434
DB 364 GCGCTGACGGGTCTGCAGCATCCGCGCGCATCTCTGATACAGGATCCATGAGCTTAAA 423

QY 1435 CAGGCCAAAGTGTGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1494
DB 424 CAGGCCAAAGTGTGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 483

QY 1495 CGAACTGTGACCTGGGGCTGACAGGCAACCATGAGTGTGTCACCCAGCTTCCGGAC 1554
DB 484 CGAACTGTGACCTGGGGCTGACAGGCAACCATGAGTGTGTCACCCAGCTTCCGGAC 543

QY 1555 AGTGCCAAAGTTTGTGAGGAGAACACCTACTGCTTCCGGCGGACCTGGAGACTGTG 1614
DB 544 AGTGCCAAAGTTTGTGAGGAGAACACCTACTGCTTCCGGCGGACCTGGAGACTGTG 603

QY 1615 CTCGGCTTTGGCAAGACCATCATGAGGAGGAGTGTGATGAGGAGGAGTGTGAGGAGTGTG 1674
DB 604 CTCGGCTTTGGCAAGACCATCATGAGGAGGAGTGTGATGAGGAGGAGTGTGAGGAGTGTG 663

QY 1675 CTCATCAACCTGATGAGCAGTGTGCTGCGGGCGGAGGCTTCCATCCGATC 1734
DB 1734 CTCATCAACCTGATGAGCAGTGTGCTGCGGGCGGAGGCTTCCATCCGATC

664 CTCATCAACCTGATGAGCAGTGTGCTGCGGGCGGAGGCTTCCATCCGATC 723
QY 1735 GGGCTCCGCAACACACGACGAGGTCTCTTGGCCAAACACCTTCTGCGTGGAGCTTAC 1794
DB 724 GGGCTCCGCAACACACGACGAGGTCTCTTGGCCAAACACCTTCTGCGTGGAGCTTAC 783
QY 1795 TTGAGAAATCTTTCAGCCTCTCTCAGCTGGAACAAGTATGCTCCAGAAAACCTAGATGAG 1854
DB 784 TTGAGAAATCTTTCAGCCTCTCTCAGCTGGAACAAGTATGCTCCAGAAAACCTAGATGAG 843
QY 1855 CAGATTAGAAGTGTCCACACAGATCCTTTCAGAAAGCAGGCTATATCTGTGCCACCT 1914
DB 844 CAGATTAGAAGTGTCCACACAGATCCTTTCAGAAAGCAGGCTATATCTGTGCCACCT 903
QY 1915 CTGGACAGGATGCTGAGGAGGAGGACAGTGTCCCTGTCTACCGCGCCCTTACCCT 1974
DB 904 CTGGACAGGATGCTGAGGAGGAGGACAGTGTCCCTGTCTACCGCGCCCTTACCCT 963
QY 1975 GGGCGGTTGCTGGATGACTGTTACTCTTTTTCAGAAAGTGTGTGGGATATACAGGTTA 2034
DB 964 GGGCGGTTGCTGGATGACTGTTACTCTTTTTCAGAAAGTGTGTGGGATATATTT 1023
QY 2035 AGCCCTTTTTCCTCCGCTGTCACCTTGAAGGTTGTGCTGCTGGCGTGGAGAGCCTCTTCC 2094
DB 1024 ACAAGTTTAAAGCCCTTTTGGTTTCCCGGTTTGGCAAGCCCTTGAAGGGGTTTNTC 1083
QY 2095 AGGTTTTCACCTGTCAGG 2111
DB 1084 CGCCCTTGGGNCCTGG 1100

RESULT 12
AAH08631
ID AAH08631 standard; cDNA; 771 BP.
XX
AC AAH08631:
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:5466.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
KW Homo sapiens.
XX
PN EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX
XX 27-AUG-1999; 99JP-0300253.
XX
XX 11-JAN-2000; 2000JP-0118776.
XX
XX 02-MAY-2000; 2000JP-0183767.
XX
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs.
XX
XX Claim 1; SEQ ID 5466; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set

comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AA892446 to AA893893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the invention of the present invention.

SQ Sequence 771 BP; 204 A; 187 C; 219 G; 158 T; 3 other;

Query Match	30.6%;	Score	750.6;	DB	22;	Length	771;
Best Local Similarity	99.0%;	Pred. No.	8, 2e-192;				
Matches	764;	Conservative	0;	Mismatches	7;	Indels	1;
Qy	21	GCTAAGNAGGGAGACTGAGGCTGAGGCTGGGGAACATCGGGCAGCATGACGCGCTCGCG	80				
Db	1	GCTAAGNAGGGAGACTGAGGCTGAGGCTGGGGAACATCGGGCAGCATGACGCGCTCGCG	60				
Qy	81	GCTCTTCTCGCACACAGGGCTGCGGGCTCGCTCGCGGGGTCTGGTGGTCTCTACCGC	140				
Db	61	GCTCTTCTCGCACACAGGGCTGCGGGCTCGCTCGCGGGGTCTGGTGGTCTCTACCGC	120				
Qy	141	GAACCGCGGCTACTTGGCAGCAGCCGCCCTGTACGAGCTTTCGCCAAGAGAGCTTTTCTC	200				
Db	121	GAACCGCGGCTACTTGGCAGCAGCCGCCCTGTACGAGCTTTCGCCAAGAGAGCTTTTCTC	180				
Qy	201	AGGCAAAATCAAGAAAGAAAGTATTTCCCATTTCCAGAAGTTAGCCAAGATGAACTTAA	260				
Db	181	AGGCAAAATCAAGAAAGAAAGTATTTCCCATTTCCAGAAGTTAGCCAAGATGAACTTAA	240				
Qy	261	TGAATCAATCAGTTCTTTGGGACCGGTGGAAAAATTTCTCAGTGAAGAGTGGACTCCCG	320				
Db	241	TGAATCAATCAGTTCTTTGGGACCGGTGGAAAAATTTCTCAGTGAAGAGTGGACTCCCG	300				
Qy	321	AAAAATTGACCAGGAAGGAAAAATCCACAGATGAACCTTTGGAGAAATTCAGAGAGCTTAG	380				
Db	301	AAAAATTGACCAGGAAGGAAAAATCCACAGATGAACCTTTGGAGAAATTCAGAGAGCTTAG	360				
Qy	381	GCTTTTGGGCTGCAAGTCCAGAAATATGTGGCCTGGGCTTCTCCAACACCATGTA	440				
Db	361	GCTTTTGGGCTGCAAGTCCAGAAATATGTGGCCTGGGCTTCTCCAACACCATGTA	420				
Qy	441	CTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCTGTGACCTGGCAGCGCA	500				
Db	421	CTCAGACTAGGGAGATCATCAGCATGGATGGTCCATCTGTGACCTGGCAGCGCA	480				
Qy	501	CAAGCTATTGGCTCAAGGGGATCATTTGGCTGGCAGTGAAGAGCAAGAACCAATA	560				
Db	481	CAAGCTATTGGCTCAAGGGGATCATTTGGCTGGCAGTGAAGAGCAAGAACCAATA	540				
Qy	561	CTTGCTTAATCTGGCTCCGGGAGCACATTCGACGCTTTCGCTCAGGAGCCAGCGAG	620				
Db	541	CTTGCTTAATCTGGCTCCGGGAGCACATTCGACGCTTTCGCTCAGGAGCCAGCGAG	600				
Qy	621	TGGAGCGATGCGCCTCAATCCGAGCAGGCCACACTAAGTGAAGACAAGAGCACTA	680				
Db	601	TGGAGCGATGCGCCTCAATCCGAGCAGGCCACACTAAGTGAAGACAAGAGCACTA	660				
Qy	681	CATCCTCAATGGCTCCAAGGCTGGATTACTAATGGAGGACTGGCCAAATATTTTACTGT	740				

PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.

PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 05-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-465566/50.
 XX P-PSDB; AAU23009.
 DR Novel polypeptides and polynucleotides useful for diagnosing,
 DR preventing, treating neural, immune system, muscular, reproductive,
 DR pulmonary, cardiovascular, renal, proliferative disorders and cancerous
 DR diseases
 XX Claim 4; SEQ ID No 105; 1180pp; English.
 PS The present invention relates to the isolation of novel human enzyme
 CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
 CC encoding them. The enzyme polypeptides of the invention may comprise the
 CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
 CC isomerases or ligases. The sequences of the invention are useful in the
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of
 CC disorders including hyperproliferative disorders (e.g. cancer),
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders
 CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
 CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders
 CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
 CC blood-related disorders (e.g. haemophilia), reproductive disorders
 CC (e.g. infertility) and infectious disorders (e.g. Influenza). The
 CC polynucleotides of the invention can also be used in gene therapy.
 CC AAS40785-AAS41684 represent cDNA sequences encoding for the novel human
 CC enzyme polypeptides of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 651 BP; 161 A; 166 C; 194 G; 130 T; 0 other;
 Query Match 26.5%; Score 650; DB 22; Length 651;
 Best Local Similarity 100.0%; Pred. No. le-164;
 Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GTGTGTGTGTCCTCGCGCGCTAAGAGGGGAGACTGAGGTGAGGTGGGGAACATCGG 61
 Db 2 GTGTGTGTGTCCTCGCGCGCTAAGAGGGGAGACTGAGGTGAGGTGGGGAACATCGG 61
 QY 62 GCAGCATGAGCGGTGCGGGCTCTTCTTCCACACCGGTGCGGGCTGTCGTCGCGGG 121
 Db 62 GCAGCATGAGCGGTGCGGGCTCTTCTTCCACACCGGTGCGGGCTGTCGTCGCGGG 121
 QY 122 GTCTGGTGTCTTACCGCGAACCAGCGGGCTTACTGCGCACCAGCCCGCTGTACGAGCTT 181

KW WW/rsp5/WMP domain; therapy; forensic; genetic mapping; diagnostic; detection; treatment; cervical; melanoma; colorectal adenocarcinoma; Wilm's tumour; retinoblastoma; sarcoma; myosarcoma; lung carcinoma; leukemia; lymphoma; dysplasia; hyperplasia; endometrium; adrenal; prostate; ss.

XX Homo sapiens.

XX WO9933982-A2.

XX 08-JUL-1999.

XX 22-DEC-1998; 98WO-US27610.

XX 21-DEC-1998; 98US-0217471.

XX 23-DEC-1997; 97US-0068755.

XX 03-APR-1998; 98US-0080664.

XX 21-OCT-1998; 98US-0105234.

XX 27-OCT-1998; 98US-0105877.

XX (CHIR) CHIRON CORP.

XX (HYSE-) HYSEQ INC.

XX Crkvenjakov R, Dickson M, Drmanac R, Drmanac S, Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA; Jones LW, Kassam A, Kennedy GC, Kita D, Labat I; Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C; Stache-Crain B, Sudduth-Klinger J, Williams LT; WPI; 1999-430243/36.

XX New isolated human polynucleotides

XX Claim 1; Page 503; 59lpp; English.

XX This invention describes novel isolated human polynucleotides obtained by screening for differential expression in colon cancer, breast cancer and lung cancer cell lines. The polynucleotides of the invention are represented in AAX98275-X99118 and encode polypeptides of protein families selected from 4 transmembrane segments integral membrane proteins, 7 transmembrane receptors, ATPases associated with various cellular activities (AAA), eukaryotic aspartyl proteases, GATA family of transcription factors, G-protein alpha subunit, phospholipase C, diacylglycerol binding proteins, protein kinase, protein phosphatase 2C, protein tyrosine phosphatase, trypsin, wnt family of developmental signalling proteins and WW/rsp5/WMP domain containing proteins. The encoded polypeptides also have a functional domain selected from Ank repeat, basic region plus leucine zipper transcription factors, bromodomain, EF-hand, SH3 domain, WD domain/G-beta repeats, zinc finger (C2H2 type), zinc finger (CCHC class), and zinc-binding metalloprotease domain. The polynucleotides encode polypeptides with similarity to known protein families and are predicted to have similar properties. The novel polynucleotides can be used to develop products for use as therapeutic agents and in forensics, genetic analysis, mapping and diagnostic applications. In particular, the product can be used for the detection and management of cancers. They can be used for treating e.g. cervical cancers, melanomas, colorectal adenocarcinomas, Wilm's tumour, sarcomas, retinoblastoma, myosarcomas, lung carcinomas, leukemias, such as chronic myelogenous leukemia, promyelocytic leukemia, monocytic leukemia, and myeloid leukemia, and lymphomas such as histiocytic lymphoma, anhydric hereditary ectodermal dysplasia, congenital alveolar dysplasia, epithelial dysplasia of the cervix, fibrous dysplasia of bone, and mammary dysplasia, hyperplasias, e.g. endometrial, adrenal, breast, prostate or thyroid hyperplasias or pseudoepitheliomatous hyperplasia of the skin.

XX Sequence 772 BP; 161 A; 207 C; 186 G; 165 T; 53 other;

Query Match 20.4%; Score 499.6; DB 20; Length 772;
Best Local Similarity 91.1%; Pred. No. 3.9e-124;
Matches 545; Conservative 0; Mismatches 49; Indels 4; Gaps 3;

QY 1098 GAAATTTGCACTGATGGCTCAGAAGGCTTACGTGAGAGATGACCTACCTCACAGC 1157

Db 83 GAAATTTGCACTGATGGCTCANAAGGCTTACGTTTGGAGAGTATGACCTACCTCACAGN 142
QY 1158 AGGGATGCTGACCAAACTGCTTCCGACTGCTCCATCGAGGACCCATGTTGAAGT 1217
Db 143 AGGGATGCTGACCAAACTGCTTCCGACTGCTCCATCGAGGACCCATGTTGAAGT 202
QY 1218 GTTCAGCTCCGAGGCGGCTGGCAGTGTGTGAGTGAGCGCTGCAGATCCTCGGGGCTT 1277
Db 203 GTTCANCTCCGAGGCGGCTGNCAGTGTGTGAGTGAGCGCTGCAGATCCTCGGGGCTN 262
QY 1278 GGGCTACACAAGGGACTATCCGTAGGAGCGCATACTGCGTGACACCGCGCATCTCTCTCAT 1337
Db 263 GGGCTACACAAGGGACTATCCGTAGGAGCGCATACTGCGTGACACCGCGCATCTCTCAT 322
QY 1338 CTTTCAGGGAACCAATGAGATTCTCCGGATGTACATGCCCTGAGGGTCTGACAGATGC 1397
Db 323 CTTTCNAGGGAACCAATGAGATTCTCCGGATGTACATGCCCTGAGGGTCTGACAGATGC 382
QY 1398 CGGGCGCATCTGACTACCAAGATCCATGAGCTTAAACAGCGCAAAAGTGACACAGTCAT 1457
Db 383 CGGGCGCATCTGACTACCAAGATCCATGAGCTTAAACAGCGCAAAAGTGACACAGTCAT 442
QY 1458 GGATACCGTTGGCGGAGGCTTTCGGGACTCCCTGGGCGCAAACTGTGGACCTGGGGCTGAC 1517
Db 443 GGATACCGTTGGCGGAGGCTTTCGGGACTTNCCTGGGCGCAAACTGTGGACCTGGGGCTGAC 502
QY 1518 AGGCAACCATGGAGTTGTGCCACCCAGTCTTGGGACAGTGCCCAACAAGTTTGAGGAGAA 1577
Db 503 AGGCAACCATGGAGTTGTGCCACCCAGTCTTGGGACAGTGCCCAACAAGTTTGAGGAGAA 562
QY 1578 CACCTACTCTTCGGC-CGGAGCGTGGAGACACTGCTGCTCGGCTTTGGC-AAGACCATC 1635
Db 563 CACCTACTCTTANCTCNGACCGTGAGACACTTGTGNTNCTTGGCAAAAGACCATC 622
QY 1636 ATGGAGGAGCAGCTGG--TACTGAAGCGGGTGGCAACATCTCATACCTGTATGG 1691
Db 623 ATGGANGANGANNNGTGTCTNAACNGNNTGGCCCAACATCTCATCAACCTGTATTTG 680

Search completed: September 6, 2003, 17:43:01
Job time : 681.218 secs .

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: September 6, 2003, 10:14:13 ; Search time 140.926 Seconds
(without alignments)
7679.711 Million cell updates/sec

Title: US-09-945-326-1
Perfect score: 2452
Sequence: 1 cgtgtgtgtctccctcgccg.....actgttaaaaaaaaaaaaaa 2452

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	168.6	6.9	1341	4	US-09-252-991A-11446
2	165.6	6.8	1212	4	US-09-252-991A-11389
3	155.6	6.3	1173	4	US-09-252-991A-9487
C 4	155.6	6.3	1209	4	US-09-252-991A-9346
C 5	143.4	5.8	4403765	3	US-09-103-840A-2
C 6	143.4	5.8	4411529	3	US-09-103-840A-1
C 7	139.4	5.7	3515	4	US-09-221-017B-824
8	134	5.5	1395	4	US-09-252-991A-16537
C 9	134	5.5	1665	4	US-09-252-991A-16061
10	132.6	5.4	33312	4	US-08-311-731A-121
11	123.2	5.0	1260	4	US-09-252-991A-14526
C 12	123.2	5.0	1260	4	US-09-252-991A-14838
13	111.2	4.5	1155	4	US-09-648-004-5
14	111.2	4.5	17417	4	US-09-648-004-27
15	105.8	4.3	1170	4	US-09-328-352-2254
16	102	4.2	1275	4	US-09-252-991A-8082
17	102	4.2	1488	4	US-09-252-991A-8036
C 18	102	4.2	2055	4	US-09-252-991A-8231
19	97	4.0	1185	4	US-09-328-352-2675
20	92.2	3.8	4403765	3	US-09-103-840A-2
21	92.2	3.8	4411529	3	US-09-103-840A-1
22	91.8	3.7	849	4	US-09-252-991A-12573
C 23	91.8	3.7	969	4	US-09-252-991A-12911
24	91.8	3.7	1257	4	US-09-252-991A-12881
25	90.4	3.7	1539	4	US-09-364-230-29
26	87.2	3.6	1164	4	US-09-328-352-2316
C 27	86.6	3.5	1440	4	US-09-252-991A-2675

28	86.6	3.5	2592	4	US-09-252-991A-3003	Sequence 3003, Ap
29	85.4	3.5	1553	4	US-09-364-230-31	Sequence 31, Appl
30	74.4	3.0	2079	4	US-09-252-991A-11426	Sequence 11426, A
31	74	3.0	1248	4	US-09-328-352-573	Sequence 573, App
32	72	2.9	1209	4	US-09-328-352-155	Sequence 155, App
33	70.8	2.9	1056	4	US-09-364-230-27	Sequence 27, Appl
34	70.2	2.9	1230	4	US-09-252-991A-14780	Sequence 14780, A
35	66.2	2.7	1182	4	US-09-252-991A-6534	Sequence 6534, Ap
C 36	66.2	2.7	1419	4	US-09-252-991A-6499	Sequence 6499, Ap
C 37	60	2.4	513	4	US-09-252-991A-9318	Sequence 9318, Ap
38	57.6	2.3	1149	4	US-09-328-352-724	Sequence 724, App
39	54.8	2.2	276	4	US-09-252-991A-9426	Sequence 9426, Ap
40	53.8	2.2	958	4	US-09-155-183-10	Sequence 10, Appl
41	53.8	2.2	1227	4	US-09-328-352-1449	Sequence 1449, Ap
42	53.4	2.2	766	4	US-09-364-230-33	Sequence 33, Appl
43	53.4	2.2	1200	4	US-09-328-352-234	Sequence 234, App
C 44	52.4	2.1	7218	1	US-08-232-463-14	Sequence 14, Appl
45	50.8	2.1	1302	4	US-09-252-991A-14660	Sequence 14660, A

ALIGNMENTS

RESULT 1

US-09-252-991A-11446/C
; Sequence 11446, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 11446

; LENGTH: 1341

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11446

Query Match	6.9%	Score 168.6;	DB 4;	Length 1341;
Best Local Similarity	49.9%	Pred. No. 2.9e-37;		
Matches	553;	Conservative 0;	Mismatches 529;	Indels 27; Gaps 4;
Qy	308	AGGTGGACTCCGAAAAATTTGACCAAGAGGGAATCCCGAGATGAACTTTGGAGAAAT	367	
Db	1220	AGCGCGCGCGCGCAGATCGACGAGCAANTCAGTTCCCGCATGGACATGTGGCGCAAGT	1161	
Qy	368	TGAAGAGCCTAGGCGCTTTTGGCTGCAAGTCCAGAGAAATATGTCGCTGGCGTTCT	427	
Db	1160	TCGCGGAGATGGCGCTGCTCGGCATCACCGTCGACGAGGATACGCGGTTCCGCGCTGG	1101	
Qy	428	CCAAACCACTGTACTCAAGACTAGGAGGATCATCAGATGG---ATGGGTCCATCAGT	484	
Db	1100	GCTACCTGGCGCGCGCGTGGTTCATGGAAGAGATCAGCGCGGCTCGGCTCGGTGGCGC	1041	
Qy	485	TGACCTGGCGCGCGCGCGCGCTATTGGCTTCAAGGGGATCATCTTGGCTGGCGCTAGG	544	
Db	1040	TCCTCTATGGCGCGCATTCACACCTCTCGGTCAACCATGATCAAGCGCAACGCGCG	981	
Qy	545	AGCAGAAAGCAATACTTGGCTTAACCTGGCGCGGAGACATTCGAGCTTCTGCG	604	
Db	980	AAACAGAGCGCGCTACCTACCTCGCGCGCTGATTCGCGGCAACACATCGGTGGCTGGCA	921	
Qy	605	TCACGAGGAGCGCGCTAGGAGGATGACGCTCAATCCGAGGAGACACACATAGTG	664	
Db	920	TGAGCAACCAACCGCGCGCTCCGAGTGGTCTCGATGAAGCTGCGCGCGCGCGGTCG	861	

```

665 AGACAAAGAGACATACATCCCTCAATGCTCAAGGCTCGGATTAATGAGACTGG 724
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
860 GCAC-----CGCTTCTGCTCAAGCGCAGCAAGATGGATCAACACGCTCCGATG 807
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
725 CCAATATTTTACTGTGTTTGAAGACATGAGTGGTGTGATTTCTGATGGATCAAGTGAAG 784
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
806 CCATACCTATGTGATCTACGCCAAGACCGACGCGGACAA-----GGCGGCGC 759
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
785 ACAAATACAGACATTCATAGTAGAAGAGACTTTGGTGGATCATACTATGGGAACCGC 844
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
758 ACGGATACCGCGCTTCATCGTCGACGCGACTGGAAGGCTTCAGTCGCGCGCGAAGC 699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
845 ARGATAATTTAGCATTTCGGGGCTCCAAACACTTGTGAAGTCCATTTTGAAGAACACCAAGA 904
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
698 TGGACARAGTGGGCTATCGCGGCTCGAACACCTGCGAATGATCTTCAGGAGCTGAGG 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
905 TACCTGTGGAACATCTCTTGGAGAGTTCGGAGATGGGTTTAAAGTGGCCATGAACATCC 964
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
638 TACCGAGGAGAACTCTCGCGCGGTCAACGCGGCGGTCAAGGTACTGATGAGCGGCC 579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
965 TCAACACGCGCGCTTCAGATGGCAGCGCTGCTGGCTGTCTCAAGAGATTTGATTG 1024
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
578 TCGACTACGAGCGGCTGGTCTCTTCGCGCGCGCGGTGGGAATCATGCGGCTGCATGG 519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1025 AAATGACTGTGATAGTACCGCTGCACAAAGAAACAGTTTAAAGAGGCTCAGTGAATTG 1084
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
518 ACGTGGTGGTCCGTACATCCATGACCGCGCGCAGTTTCGGCCAGAGCATCGCGGAATCC 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1085 GATTGATCAGGAGAAATTTGCATGATGCTCAGAGGCTTACGATGAGAGATGA 1144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
458 AGCTGGTCCAGGCAAGGTGGCGGACAT-----GTACCGCGCTCAACGCCACGCGCG 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1145 CCTACCTCAGCAGGAGTGTGGACCAACCTGGCTTTCCCGACTGTCTCATCGAGCGAG 1204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
404 CCTACCTGTACCGCTGGCGCGCGCTCGACCGCGCGAGACCGCCGCAAGAGCGCG 345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1205 CCATGGTGAAGTGTTCAGTTCGAGCGCGCTGGAGTGTGTGAGTGAAGCGCTCGAGA 1264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
344 CCGGGGTGATCCTCTACAGCGCGGAGCGCGCCACCCAGATGGCCCTGGAGCGCATCAGA 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1265 TCCTCGGGGCTTGGCTTACAAAGGACTATCCGTACGAGCGCATACTCGCTGACACCC 1324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
284 TCCTCGGGCGCAAGCTTACATCAACGAGTTTCCACCGCGCGCTGTTCGCGAGCGCCA 225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1325 GCATCCTCTCATCTTCGAGGGAACCAATGAGATTCCTCGGATGTACATCGCCCTGACGG 1384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
224 AGCTTACAGATCGCGCGCGCACCGAGATCCGCGCATCTGATCGCGCGCAGC 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1385 GTCTGCACATGCGCGCGCATCTGACT 1413
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
164 TGTTCACGAACCGCTGAAGCTTCGCT 136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 2
US-09-252-991A-11389
; Sequence 11389, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11389
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa

```

```

US-09-252-991A-11389
Query Match      6.8%; Score 165.6; DB 4; Length 1212;
Best Local Similarity 50.1%; Pred. No. 1.9e-36;
Matches 538; Conservative 0; Mismatches 509; Indels 27; Gaps 4;

QY 308 AGGTGGACTCCCGAAAAATTTGACCAGGAAGGAAATCCAGATGAACATTTTGAGAAAT 367
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 137 AGCGCGCGCGCGCAGATCGACGAGCAATCAGTTCCCCCATGGACATGTGGCGCAATG 196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 368 TGAAGAGCTAGGGCTTTTGGGCTGCAAGTCCAGAGATATATGTTGGCTGGGCTTCT 427
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 197 TCGGGAGATGGCTGCTCGCATCACCGTTCGAGGAGTACGGCGTTCGCGCGCTGG 256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 428 CCAACACCATCTACTCAAGACTAGGGAGATCATCAGCATGG---ATGGGTCCATCATG 484
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 257 GCTACTGGCCACACCGCTGTGATGGAAGAGATCAGCGGGCGCTCGGCTCGGTGGCGC 316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 485 TGACCTTGGCAGCGCACCGAGGCTATTGGCTTCAAGGGGATCATCTTGGCTGGCACTGAGG 544
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 317 TCTCTATGGCGGCTTCCAACTCTCGGTCAACCATCAAGCGCAACGCAACGCGC 376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 545 AGCAGAAAGCCAAATTTACTTGGCTTAACTTGGGCTCGGGGAGACATATTCAGCCCTTCTGCC 604
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 377 AACAGAAGCGCGCTACCTGCGCGCTTGTATCCGCGCAACACATCGGTGGCGGA 436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 605 TCACGAGCCACCCAGTGGAGCGATGCGGCTCAATCCGGAGCAGAGCCACACTTAAGTG 664
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 437 TGAGCGAACCCAAACCGCGCTCCGACGTGCTCGATGAAGCTGCGCGGACCGGCTCG 496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 665 AAGACAAGAACACTACATCTCAATGGCTCCAAGTCTGGATTTACTTAATGAGACTGG 724
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 497 GCGAC-----CGCTTCTTCTCAACGCGCAGCAAGTGTGATCAACACGCTCCCGATG 550
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 725 CCAATATTTTACTGTGTTTGAAGACTGAGTGTGATTTCTGATGATGATGATGAAAG 784
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 551 CCATACCTATGTGATCTACGCCAAGCGCGCGGACAA-----GGCGCGC 598
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 785 ACAAATACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGAAACCGG 844
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 599 ACGGATCACCGGTTTCATCTCGCGCGGCTGGAAGGCTTCACTCGCGCGCGGCAAGC 658
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 845 AAGATAAATTAGCATTCGGGGCTCCAACACTTGTGAAGTCCATTTTGAAGAACCAACA 904
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 659 TGGACAAGCTGGCATCGCGCTCGAACACCTCGCACTGATCTTCCAGGAGCTGCGAG 718
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 905 TACCTGTGAAACATCTTGGAGAGTTCGAGATGGGTTTAAAGTGGCCATCAACATCC 964
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 719 TACCCGAGGAGAACGCTCTCGCGCGGTCAACGCGGGTCAAGGTACTGATGAGCGGCC 778
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 965 TCAACAGCGCGCGTTTACAGATGGGAGCGTCTGGCTGGCTGCTCAAGAGATTCATG 1024
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 779 TCGACTACGAGCGGTGTCTTTCGCGCGCGCGGTAATCATGCGAGCGGTGATGG 838
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1025 AATGACTGCTGAGTACGCTGCACAAAGAAACAGTTTAAACAGAGGCTCAGTGAATTTG 1084
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 839 ACGTGGTGGTGCCTACATCCATGACCGCGCGCAGTTCGCGCAGACATCGCGGAATTC 898
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1085 GATTGATTCAGGAGAAATTTGCACTGATGGCTCAGAAAGGCTTACGTCATGAGAGATGA 1144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 899 AGCTGGTCCAGGCAAGGTGGCGGACAT-----GTACACCGCGCTCAACGCCACCGCGG 952
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1145 CCTACTCTCAGCAGGATGCTTGACCAACCTGGCTTCCGACTGCTCCATCGAGCGAG 1204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 953 CCTACTGTACGCGGTGGCGCGCTTCGACCGCGCGAGACGACCCGCAAGGACCGG 1012
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1205 CCATGTTGAAGTGTTCAGCTCCGAGCGCGCTGGCAGTGTGTGAGTGAAGCGCTCGAGA 1264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1013 CCGGGGTGATCCTTACAGCGCGGAGCGCGCCACCCAGATGGCGCTTGGACGATCCAGA 1072
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1265 TCCTCGGGGCTTGGGCTACACAGGAGACTATCGGTACGAGCGCATCTCGGTGACACCC 1324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1073 TCCTCGGGGCAACGGCTACATCAACAGAGTTTCCACCGCGCGCTTTCGCGAGCGCCA 1132
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Db 3651995 GCGCCGAACGCGGTGAGCGGATCTGGGCTTCAITTCGGCGGCTCGAAGTGTCTCGCCT 3651936
QY 1226 CCGAGCGCCCTGGCGAGTGTGTGAGTGGCGCTGCAGATCCTCGGGGGCTTGGGCTACA 1285
Db 3651935 CCGAGCTCGCATGGAGGTACACCGATGCGGTACAACTGTTCGGGGGCGCGCTACA 3651876
QY 1286 CAGGGGACTATCCGTACGAGCGCATCTGCGTGACACCGCATCTCTCTCATCTTCGAGG 1345
Db 3651875 CCACGACTTCCGCTCGAGGGTTTATGCGCGAGCGCAAGATCACCCATCTATGAGG 3651816
QY 1346 GAACCAATGAGATCTCCCGATGTACATCGCCCTGACGGGTCTGC 1390
Db 3651815 GCACCAATCAGATTACGCGGTAGTGATGTGCGGGGCGCTACTGC 3651771

RESULT 6
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103, 840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv

US-09-103-840A-1

Query Match 5.8%; Score 143.4; DB 3; Length 4411529;
Best Local Similarity 48.2%; Pred. No. 3.3e-28;
Matches 513; Conservative 0; Mismatches 531; Indels 21; Gaps 3;

QY 326 TTGACGAGGAGGAAATCCAGATGAACTTTTGAGAAATTTGAAGAGCTAGGCTTTT 385
Db 3657964 TGCACGAGAAGCGCGGTTCGCCGAAGAGCGCTAGTGGCGTCAATTCCTCCGGTTTCA 3657905
QY 386 TTGGGCTCAACTCCAGNAGATATGTGGGCTTCTCCCAACACCATGACTCAA 445
Db 3657904 ACGCCGTCCACATTCGGGAGGATACGGCGTCAAGGTGCGGACTCGGTAGCTACCTCA 3657845
QY 446 GACTAGGGGAGATCATCAGCATGGATGGTCCATCATCTGTACCCCTGCACCGCACAGG 505
Db 3657844 TCGTGATCGAAGAGGTGCGCGCTGACGCTGCTGATCCCGCGGTCAACA 3657785
QY 506 CTATTGGGCTCAAGGGATCATCTTTGGCTGGCACTGAGGAGCAGAAAGCCAAATCTGC 565
Db 3657784 AGCTGGGACCATTCGGCGCTGATCTCTCGGGGCTCCGAGGAGCTGAAGACAGGTGCTGC 3657725
QY 566 CTAAACTGGCGTCCGGGAGCACATTCAGCGTTCTGCCTCAGCGAGCCAGCCAGTGGGA 625
Db 3657724 CGCGGTTGGCGCGGAGGCGGATGGCTTCCTATGATTTAGTGAACGAGCGCGGCA 3657665
QY 626 GCGATGACGCTCAATCCGAGCAGAGCCACACTAAGTGAAGACAAAGCAAGCACTACATCC 685
Db 3657664 CTGACGCGGCTCCATCGGACGCGGCGCC-----AAGCGGATGGGATCACTGGATTC 3657611
QY 686 TCAATGGCTCAAGGTCTGGATTACTAATGAGGACTGGCCAATATTTTACTGTGTTTG 745
Db 3657610 TCACGGCGGCAAGTGTGGATCACCAACGCGGCAAGTCGACCTGGTACACCGGTATGG 3657551
QY 746 CAAAGACTGAGGTGCTTGTATTCATGATGATCAGTGAAGACAAATCAGCAATTCATAG 805
Db 3657550 CGGTGACCGA-----TCCCGACCGGGCGGCGCAACGCGATCTCGGGGCTTCATGG 3657503

QY 806 TAGAAAGAGACTTTGGTGGAGTCACTAATTTGGGAAACCCCAAGATAAATTTAGGCATTCGGG 865
Db 3657502 TGCACAAAGACGACGAGGGGTTTCCCGTGGTCCGAAAGAGCGAAGTTCGGGATCAAGG 3657443
QY 866 GTTCCAAACACTTTGTGAAGTCCATTTTGAAGAACCAACCAAGTACTCTGTGAAACAACTTCCTTTG 925
Db 3657442 GGTACCCGACCCAGCGAGTGTACTTCGAGAAGTCCCGGATCCCGGCGCATCGCATCATCG 3657383
QY 926 GAGAGTCGAGATGGGTTTAAAGTGGCCATGAACATCCTCAACAGCGCGCGGTTCAGCA 985
Db 3657382 GTGACCCGCTACCGGTTTCAAGACCGCGCTGGCCACCTTGGACACACCCGTCGCCAGA 3657323
QY 986 TGGGAGCGCTCGTGGCTGGCTCTCAAGAGATTGATTGAATGACTCTGAGTACGCT 1045
Db 3657322 TTGGCGCGCAGCGCGCTGGTATCGCCAGCGCGCTGGATGCTGCCATCCCTACACCA 3657263
QY 1046 GCACAAGGAAACAGTTTAAAGAGAGGCTCACTGAATTTGGATTGATTTCAGGAGAAATTTG 1105
Db 3657262 AGGACCGCAAGCAGTTTCGGTGCATCGATCAGTACGTTTCAGGCGGTTCAGTTCATGCTGG 3657203
QY 1106 CACTGATGGCTCAGAAGGCTTACGTCATGGAGACTATGACTACCTACCTCACAGCAGGATGC 1165
Db 3657202 CCGACATGGCATGAAGGTGGAGCGGCTCGGCTGATGGTTTACTCCGCGCGCGCC---C 3657146
QY 1166 TGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGAGGAGCATGGTGAAGGTGTTCAGCT 1225
Db 3657145 CGCGCAAGCGGTGAGCGGATCTGGCTTCATTTTCGGCGGCTCGAAGTGTTCGCT 3657086
QY 1226 CCGAGGCGCGCTGGCAGTGTGTGAGTGGCGGCTGAGAGTCCCTCGGGGGGTGGGTGATCA 1285
Db 3657085 CCGAGTTCGGATGGAGGTCAACCATGCGGTACAACTGTTTCGGCGCGCGGCTGATCA 3657026
QY 1286 CAAGGACTATCCGTACGAGCGCATCTGCGGTGACACCCCGCATCTCTCTCATCTTCGAGG 1345
Db 3657025 CCACGACTTCCCGTTCGAGCGGTTTATGCGCGACCGCAAGATCACCCAGATCATGAGG 3656966
QY 1346 GAACCAATGAGATCTCCGATGATGATGTCGCGGGGTAGTGTGTCGCGGGCGCTACTGC 3656921
Db 3656965 GCACCAATGAGATTCAGCGGTAGTGTGTCGCGGGCGCTACTGC 3656921

RESULT 7

US-09-221-017B-824/c
; Sequence 824, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Pp1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Pp1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:

Tue Sep 9 10:14:36 2003

APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Monroy, Gladys H.
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO.: 824:
SEQUENCE CHARACTERISTICS:
LENGTH: 3515 base pairs
TYPE: nucleic acid
STRAINEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE: PORYPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1...3515
1S-09-221-017B-824

Query Match	5.7%	Score 139.4	DB 4	Length 3515
Best local Similarity	48.7%	Pred. No. 7.9e-29		
Matches	516	Mismatches 516	Indels 27	Gaps 4
	Conservative			
Qy	321	AAAAATTGACCGAAGGAAATCCGACATGAACTTTGGAGAAATTTGAAGAGCGCTAGG	380	
Db	1655	AGAAATTTGACGACGGAAGCTTTTCCCATCGAAACAGATAAGAGATGGCGGAATCGG	1596	
Qy	381	GCTTTTGGGCTGCAAGTCCGACAGAAATGATGGCCCTGGGCTTCTCCAACACCATGTA	440	
Db	1595	ACTGATGGGATCCCTTTCCCGCTTGAAATAGGTTGGTCCGGCGGTACAAACGTAATGTA	1536	
Qy	441	CTCAAGACTAGGGAGATCATCAGCATGGATGGGTCCATCACT--GTGACCTGGCAGC	497	
Db	1535	TTCAATGGCAGTGGAGAACTAAGCCGCTGATGTGTACCAACAGGGCGTGGTCTCTCTGC	1476	
Qy	498	GCACGAGCTATTGGCCTCAAGGGGATCATCTTTGGTGGCACCTGAGGAGCAGAAAGCCAA	557	
Db	1475	ACACACATCTTTGTGCGCAGCTCCCATCTACGAAATTTGGAACAGAAACAAAGATGAA	1416	
Qy	558	ATACTTGGCTAAATGGCGTCCGGGAGCACATTCAGCCCTCTGGCTCACGAGCGCAGC	617	
Db	1415	GTATTTCCGAAACTCTGTCTGCGAATGGATCGGTGCAATTTGGACTTACCGAACCCAA	1356	
Qy	618	CAGTGGGAGCGATCAGCGCTCAATCCGAGCAGAGCCACACTAAGTGAAGACAAAGACA	677	
Db	1355	TGCAGGTACGATGCTGCTGCTCAGCAAACTTTTGCCTGAAGAAAGGAAG-----ATCA	1302	
Qy	678	CTACATCTCAATGGCTCCAGGTCTGGATTACTAATGGAGGACTGGCCAAATATTTTAC	737	
Db	1301	TTTCGTACTTAACGGTACAAGATTTTATCCCAATGCCGAATATGCTCAGGTATACGT	1242	
Qy	738	TGTGTTTCCAAGACTAGGTCGTTGATCTGTGATGATCAGTGAAGACAAATATCAGC	797	
Db	1241	GSTATTTGCCATGACAGA-----TAAGAGCCAGGGCACAGCTGGTATCACCGC	1194	
Qy	798	ATTCATAGTAAAGAGACTTTGGTGGAGTCACATAATGGGAAACCCGAAAGATAAATTAGG	857	
Db	1193	TTTATCGTTGAAAAGAACTCCCGGTTCTCTATCGAAGAAAGAGCTTAAGATGGG	1134	
Qy	858	CATTCGGGCTCCACACTGTGAAGTCCATTTTGAACACCAAGATACCTGTGGAAA	917	
Db	1133	TATCCGTGGTCCGGCTACATCGCACTTATTTTGAACACTGTATCTGCTGCTAAAGAGAA	1074	

RESULT 8

```

RESULI 8
US-09-252-991A-16537
; Sequence 16537, Application US/09252991A
; Patent NO. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16537
; LENGTH: 1395
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16537

```

Query Match 5.5%; Score 134; DB 4; Length 1395;
Best Local Similarity 48.8%; Pred. No. 1.5e-27;
Matches 430; Conservative 0; Mismatches 440; Indels 12; Gaps 2;

[illegible]

Db 686 GTAGCAACAGCTTCATCAGCCACGCGCGGCACGCGGACTTCGCCATCGTCTTCGCCGTCA 745
Qy 752 CTGAGGTCTGTGATTCTCATGATCAGTGAAGACAAATACACAGCATTCATAGTAGAAA 811
Db 746 CCGACAGTACGACACACAGCGCGCAAGCCGCAACGCGGTGACGCGCTTCTCTGTGGACA 805
Qy 812 CAGACTTTGGTGGAGTCACTAATGGGAACCCGGAAGATAAATTAGGCATTCGGGGCTCCA 871
Db 806 AGGCGACGCGCGGATACCGTCCGCGCGCGCGGCGGAATGCGTGAGCAACCGTGGCTACC 865
Qy 872 ACATCTGTGAAGTCCATTTTCAACACACCAAGATACCTGTGGAAACATCCTTTGGAGAGG 931
Db 866 ACACCTAGGAGATCTCTTGACGATTGCGCGGTGCGCGCTCCAAAGTGTCTCGCGGAGG 925
Qy 932 TCGGAGATGGGTTTAAAGTGGCCATGAACATCCTCAACAGCGCGGCTTCAGCATGGGA 991
Db 926 TCGGAAGGGTGGGAAGTGGCCAAACGCTGGCTACCGCGCGGACGGGTGATGGTCGCG 985
Qy 992 GCGTCTGGTGGCTCTCAAGAGATTGATTGAATGACTGCTGAGTAGCTGCTGCACAA 1051
Db 986 CCAACTGTGCGGCCAGGCGCGCGCGCGCTGACCTGCTGCGTGGCGCGCGGATC 1045
Qy 1052 GGAACAGTTTAAACAGAGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCACTGA 1111
Db 1046 GCAAGCAGTTTCGGCAGCGCATCGGAGCTACCGAGTGGTTCCTTCAAGCTCGCGGACA 1105
Qy 1112 TGCGTCAAGAGCTTACGTCATGGAGAGTATGACCTACCTCACAGCAGGATGCTGGACC 1171
Db 1106 TGCGCAGCAGATCTCGCGCGAGCTGATGACCTGCACACCGCTGGAAGATGGACC 1165
Qy 1172 AACCTGGCTTTCGCGACTGCTCCATCGAGCGCCATGCTGGGGCTTGGGCTACACAGGG 1291
Db 1166 AGGCGACCATGACCGACGG-----CGAGCGCGCATGCGCAAGCTGTTTCGCCAGCGAGA 1219
Qy 1232 CCGCCTGGCAGTGTGTAGTAGGCGCTGCAGATCCTCGGGGCTTGGGCTACACAGGG 1291
Db 1220 CCGTCGGCAAGTTCGCGCAGAGGCGTGCAGATCTTCGGCGCATGGGCTGATGGATG 1279
Qy 1292 ACTATCCGTACGAGCGCATCTGCTGACACCGCATCCTCTCATCTTCGAGGGAACCA 1351
Db 1280 AAGGACCGTTCGAGCGCATCTGGCGCAACGCGCGATGCAAGGATCGGGAGGCACTT 1339
Qy 1352 ATGAGATTCCTCGGATGTACATCGCCCTGACGGGTCTGCAGC 1393
Db 1340 CGGAATCCAGCGGCACATCGTTTCCCGGAACTCTGCGGC 1381

RESULT 9

US-09-252-991A-16061/c
; Sequence 16061, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16061
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16061

Query Match 5.5%; Score 134; DB 4; Length 1665;
Best Local Similarity 48.8%; Pred. No. 1.7e-27;
Matches 430; Conservative 0; Mismatches 440; Indels 12; Gaps 2;

Qy 512 GCTCAAGGGGATCATCTTGGCTGGCAGTGAAGAGCAAAAGCCAAATATCTTGCCTAAAC 571
Db 1364 GGCCTTCGAAGATCTCTATGGCTTCACGCGGAGAACAGCTCGGCACATACCTTGTGGCCT 1305
Qy 572 TGGCGTCCGGGAGACATATGAGCCTTCTGCTCACGAGAGCCAGCCAGTGGAGCGATG 631
Db 1304 GCGTGAAGGGGAAAAGACCGATGCTTTCGCCCTCACGAGCGCGCGCTCCGAGC 1245
Qy 632 CAGCCTCAATCCGGAGAGAGCCACACTAAGTGAAGACAAAGACACTACATCTCTCAATG 691
Db 1244 CCAATTCGATCAAGACCCGCGGTGCGGAGCGGCGC-----CTTCGTGATCAACG 1191
Qy 692 GCTCAAGGCTGAGATTACTAATGAGGAGCTGGCCAATATTTTACTGTGTTTGAAGA 751
Db 1190 GTAGAACGACTTCATCAGCCACGCGGCGACGCGGACTTCGCCATCTCTTCGCCGTCA 1131
Qy 752 CTGAGGTCGTGATTCTGATGATCAGTGAAGACAAATACAGCAGTTCATAGTAGAAA 811
Db 1130 CGCAGACTAGAGCACACCGGCCAAGCGCAACGCGGTGACGCGCTTCTGTTGGACA 1071
Qy 812 GAGACTTTGGTGGATCACTAATGGGAACCCGAGATAAATTAGGCATTCGGGCTCCA 871
Db 1070 AGGCGACCGCGGATGACCTGCGCGCGCGCGCGGCGGCGGCGCTGCTTACC 1011
Qy 872 ACACCTTGTGAAGTCCATTTTCAAAACACCAAGATACCTGTGGAAGAACATCCTTGGAGAG 931
Db 1010 ACACCTAGAGATCTTCTTCGACGATTTCCCGGCTGCGGCGCTTCCAAAGTGTCTCGCGGAG 951
Qy 932 TCGGAGATGGGTTTAAAGTGGCCATGAACATCCTCAACAGCGCGGCTTACGATGGGA 991
Db 950 TCGCAAGGGCTGGGAAGTGGCCAAACGCGTGCCTACCGCGGAGCGGTGATGTCGCCG 891
Qy 992 GCGTGGTGGCTGCTGCTCAAGAGATTGATTGAATGACTGTGAGTACGCTGCACAA 1051
Db 890 CCAACTGTGTCGGCAGCGCGCGCGCTCGACCTGCTGCTGCGCTGGCGCGCGGATC 831
Qy 1052 GGAACAGTTTAAACAGAGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCATGA 1111
Db 830 GCAAGAGTTTCGGCCACCGCATCGGAGCTACGAGCGGCTTCTTCAAGCTCGCGGACA 771
Qy 1112 TGGCTCAGAAGGCTTACGTCATGAGAGTATGACCTACCTACACAGGAGGATCTGGACC 1171
Db 770 TGGCAGCAGATCCGTGCGCGGAGCTGATGACCTGACACCGCTGGAAGATGGACC 711
Qy 1172 AACCTGGCTTTCGCGACTGCTCCATCGAGCAGCCATGGTGAAGTGTTCAGTCCGAGG 1231
Db 710 AGGCGACCATGACCGACGG-----CGAGCGCGCATGCGCAAGCTGTTTCGCCAGCGAGA 657
Qy 1232 CCGCCTGGCAGTGTGAGTGAAGGCGCTGCAGATCTCGGGGCTTGGGCTACACAGGG 1291
Db 656 CCGTCGCAAGGTCGCGGAGAGCGGTGCAGATCTTCGCGCGCATGGGCTGATGGATG 597
Qy 1292 ACTATCGTACGAGCGCATCTCGGTGACACCGCATCTCTCTCATCTTCGAGGGAACCA 1351
Db 596 AAGACCGGTTCGAGCGCATCTGCGCGCAACGCGCGGATCGAAGGATCTCGGAGGCGACTT 537
Qy 1352 ATGAGATTCCTCGGATGTATCATCGCCTGACGGGTCTGCAGC 1393
Db 536 CGGAATCCAGCGGCACATCGTTTCCCGGAACTGCTGCGGC 495

RESULT 10

US-08-311-731A-121
; Sequence 121, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:

ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/311,731A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: C0044/7125
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/720-3500
TELEFAX: 617/720-2441
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 3312 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: MYCOBACTERIUM LEPRAE

US-08-311-731A-121

Query Match 5.4%; Score 132.6; DB 4; Length 33312;
Best Local Similarity 47.4%; Pred. No. 2.3e-26;
Matches 510; Conservative 0; Mismatches 544; Indels 21; Gaps 3;
QY 326 TTGACCGAGGAGAAATCCAGATGAAATTTTGAGAAATTAAGAGAGCTAGGGCTTT 385
DB 24500 TGGACCGAGCTGCTCGATTTCCCGAGGAGCGCTGGCAGCCCTGAATGCATCAGTTTCA 24559
QY 386 TTGGCGCTCAAGTCCCAAGAAATATGTTGGCCCTGGGCTTCTCCAAACCATGACTCAA 445
DB 24560 ACGTATCCACGTTCCCGAGGAGTATGTTGTCAGGTGGGATTCGGTAGCGGCTTGA 24619
QY 446 GACTAGGGAGATCATCAGCATGATGGTCCATCATCTGTGACCTTGGCAGCGCACCGAG 505
DB 24620 TTGTGATCGAAGAGTGGCGGTGTCGATGCTTCTGATCGTTGATTCCTGCACTTAACA 24679
QY 506 CTATTGGCCTCAAGGGGATCATCTTTGGCTGGCACTGAGGACGAGAAAGCCAAATACATTGC 565
DB 24680 AGCTTGGACCATGGACTCATCTTCCGCGCTTCGGAAGAGCTCAAGAAACAGGTTCTGC 24739
QY 566 CTAAACTGGCTGCGGGAGACATGTCAGCTTCTGCTCAGCGGACCGCCAGTGGGA 625
DB 24740 CATCTGTTGGCTGCGGAGGGGCGATGGCGTCTATGCAATTAAGTGAGCGGCAACCGGCA 24799
QY 626 GCGATGAGCCTCAATCCGAGAGAGCCACACTAAGTGAAGACAAAGACACTACATCC 685
DB 24800 GTGACGCTGCTGATGCGGACCGCGGCGCA-----GCTGACGGGATGACTGGATTC 24853
QY 686 TCAATGGCTCAAGTCTGGATTACTAATGAGGAGTGGCCCAATATTTTACTGTGTTG 745
DB 24854 TCAATGGCTCAAGTCTGGATTACCAACGGTGGCAAGTTCGACCTGGTACACGGTTATGG 24913
QY 746 CAAGACTAGTCTGTTGATCTGATGATGATCAAGAAAGACAAATACAGCAATCATAG 805
DB 24914 CGGTGACCATCCGACAGAGGCGCA-----ACGGATCTCGGCTTCAICG 24961
QY 806 TAGAAGAGACTTTGGTGAGTCACTAATGGGAAACCCGAGATAAATAGGCATTCGGG 865

DB 24962 TGCACAAGGACGATGAGGGATTTCAGCATTTGGCCCGAAAGAAAGAAAGAAAGCTCGGGATCAAGG 25021
QY 866 GCTCCAAACACTTTGAAGTCCATTTTGAACAAACCAAGATACCTGTGGAACAAACATCCTTG 925
DB 25022 GGTCAACCAACCAACCAACTTACTTCGATTAATTCGCATCCCGGGTATCGCATCATG 25081
QY 926 GAGAGTCTGGAGATGGGTTTAAAGTGGCCATGAACATCCTCAACAGCGGCGGTTTCAGCA 985
DB 25082 GTGAGCCGGTACTGGCTTTAAGACAGCGCTAGCAGGTTGGATTCACACGCGTCCCAACA 25141
QY 986 TGGGACGCTGCTGGTGGCTGCTCAGAGATTAATTGAATGACTGCTGAGTACGCT 1045
DB 25142 TTGGTCCCAAGCGGTGGGCTTGGCAGGCGGCTGGAGCTGCCATCGTTTATACCA 25201
QY 1046 GCACAAGGAAACAGTTTAAAGAGGCTCAGTGAATTTGGATTTGATTCAGAGAAATTTG 1105
DB 25202 AGGACCGCAAGCAATTCGCGAGTCGATTAGCACTTTCCAGTCCATTCAGTTCATGCTCG 25261
QY 1106 CACTGATGCTCAGAAAGCTTACGTCATGGAGATGATGACCTACCTACACAGCAGGATGC 1165
DB 25262 CCGATCATGGGATGAAGTGGAGGCTGCACGGTTAAATTGTCTA---CGCTGCGCTGCC 25318
QY 1166 TGGACCAACTGGCTTTCCCGACTGCTCCATCGAGGCGCATGGTGAAGGTGTTCACT 1225
DB 25319 GTGCTGAACGCGGTGAGCGGATCTGGCTTTATTTCAGCGGCGCTCGAAATGCTTGTCT 25378
QY 1226 CCGAGCGGCTGGCAGTGTGTGAGTGGCGCTGCAGATCCTCGGGGCTTGGGCTACA 1285
DB 25379 CCGACATTTGGATGGAGGTACACCGGCGCTGTGCAATTTGTTGGCGGCGGCTACA 25438
QY 1286 CAAGGACTATCGTACGAGCGCATACTGCGTGACACCGCGCATCTCTCTCATCTTCGAGG 1345
DB 25439 CTTCCGACTTCCCGTCCGAGCGGTTTCATGCGGCGGCAAGATCACAGATCTATGAGG 25498
QY 1346 GAACCAATGAGATTTCTCGGATGTATATCGCCCTGACGGGTCTGCAGATGCGCG 1400
DB 25499 GGACCAATCAGATTCAGCGTGTGGTATGTGTCGCGGCGCTGTCGCGTATCCGG 25553
RESULT 11
US-09-252-991A-14526
; Sequence 14526, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14526
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-14526
Query Match 5.0%; Score 123.2; DB 4; Length 1260;
Best Local Similarity 47.8%; Pred. No. 1.6e-24;
Matches 504; Conservative 0; Mismatches 523; Indels 27; Gaps 4;
QY 328 GACCAGCAAGGAAATCCAGATGAACTTTGAGAAATTTGAGAGCTTAGGAGCTTTT 387
DB 220 GACCGGAGCACCGTTTCCCGGCGGAGCAATCCGCGAGATGCGCACCTCGGCTTCTC 279
QY 388 GGGCTGCAAGTCCCAAGAAATATGTTGGCTTGTCCACACCATGTAATAA-- 445
DB 280 GGCATGTGTTGCCCGAGGAATGGCGGCGGCGAGACCGGCGGCACTTGGCCTACGCCATG 339

QY 1165 CTGACACACCTGGCTTCCGACTCTCCATCGAGGAGCCATGGTGAAGTGTTCAGC 1224
 Db 225 CGCGAA-----GCCGGCTTCCGCTGCTGAGCGGAGCTCGATGCCAAGCTGTTCGC 172
 QY 1225 TCCGAGCCGCTGGCAGTGTGTGAGTGGAGCGTGCAGATCCTCGGGGGCTTGGGCTAC 1284
 Db 171 TCGAGATGCCGAGGAAGTTTCTCCGCGGCGATCCAGACCTCGTGGCTATGGCTAC 112
 QY 1285 ACAAGGAGCTATCCGTACGACGCTATCTGCGTGCAGACCCGATCCTCTCATCTCGAG 1344
 Db 111 CTCAGGAGCTTCCCGGTGGAACGATCTATCGGAGCGTGGGCTCTGCCAGATCTACGAG 52
 QY 1345 GGAACCAATGAGATTCCTCGGATGATACATCGCCC 1378
 Db 51 GGCACGAGCGAGTGCAGCGCTGGTGTATCGCC 18

RESULT 13

US-09-648-004-5
 ; Sequence 5, Application US/09648004
 ; Patent No. 6498242
 ; GENERAL INFORMATION:
 ; APPLICANT: CHEN, QIONG
 ; APPLICANT: THOMAS, STUART
 ; APPLICANT: NAGARAJAN, VASANTHA
 ; TITLE OF INVENTION: BIOLOGICAL METHOD FOR THE PRODUCTION OF ADIPIC ACID AND
 ; TITLE OF INVENTION: INTERMEDIATES
 ; FILE REFERENCE: CL-1341-A
 ; CURRENT APPLICATION NUMBER: US/09/648,004
 ; CURRENT FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/252,553
 ; PRIOR FILING DATE: 1999-02-19
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 5
 ; LENGTH: 1155
 ; TYPE: DNA
 ; ORGANISM: Acinetobacter sp.
 US-09-648-004-5

Query Match 4.5%; Score 111.2; DB 4; Length 1155;
 Best Local Similarity 47.8%; Pred. No. 3.5e-21;
 Matches 499; Conservative 0; Mismatches 518; Indels 27; Gaps 5;

QY 341 AAATCCAGATGAACCTTTGGAGAAATTAAGAGCCTAGGCGCTTTGGGTGCAAGTCC 400
 Db 104 AAATCCAGCTGAATCGTGCAGCAATGAAGAACTGGGTCTTTTGGTCTCACCATT 163
 QY 401 CAGAGATATGGTGGCTGGCTTCTCCA---ACACCATGTACTCAAGACTAGGGAGA 457
 Db 164 CTGAGGAATATGAGGCTTGGCCCTGACCATGGAGAGAGGTTTACATTGCATTTGAC 223
 QY 458 TCATCAGCATGGATGGTCCATCACTGTGACCCCTGGCAGCGCACCAGGCTATTGGCCCTCA 517
 Db 224 TGGAGCTACCTCTCTGCTTTCCGTTCACTGCGCACTAACAATGGATCGGTTCAT 283
 QY 518 AGGGATCATCTTGGCTGGCTGAGGAGCAGAAACCAATATCTTGGCTTAACTGGGT 577
 Db 284 CAGGCTTAATATTATGATGGCTCCGAGAGCAGAAACAGTATTTTGGCCAGCTCTGGCAA 343
 QY 578 CCGGGAGCATATTGAGCTTCTGCTCAGCGAGCCAGCCAGTGGAGCGATGACGCT 637
 Db 344 GTGGTGAATATTGGTTTCATCTGTTTAACTGAACCTGATTCGCGTTCAGATGCTGCT 403
 QY 638 CAATCCGGAGCAGACCACTAAGTGAAGACAAAGACACTACATCTCAATGCTCA 697
 Db 404 CTTT-----AAAAACACAGCGGTGAAGAGTGGTATCATTTACATTTTAAATGGCACTA 457
 QY 698 AGGTCTGGATTACTATGAGGAGCTGCCCAATATTTTACTGTGTTTGAAGAACTGAGG 757
 Db 458 AGCGTTACATCAACAATGCAACCGCATGCGGGTGTCTTTACTGTCTATGCGCACGTAC----- 512

QY 758 TCGTTGATCTGATGGATCAGTGAACACAAAATCAGCATTCATAGTAAGAGAGACT 817
 Db 513 ---CAGTACCGAAATTAAGAGTACAGGTGGAATTTTCAGCCTTTATCGTGGACAGTAAAA 568
 QY 818 TTGCTGAGTCACTAATGGGAAACCCGAAGATAAATTAGGCATTCGGGGCTTCCACACTT 877
 Db 569 CTCCTGGTATTTCTTGGTAAAGCTGATGAAGAGATGGCCAAAGGTGCACATACCT 628
 QY 878 GTGAAGTCCATTTTGAACACCAAGATACCTG---TGGAAACATCCTTTGGAGAGTGC 934
 Db 629 GTGATGTATTTTGAACACTGTCTATTCCTGCTATTCGCTATTCGCTATTCGCTATTCGCT 688
 QY 935 GAGATGGTAAAGTGGCCATGAACATCTCAACAGCGCGCTTCAGCATGGGCGACG 994
 Db 689 GTGTAGGTTTAAAACTGCAATGAAGTACTTGTAAAGCGCTATTTCATATTCTGCTCAT 748
 QY 995 TCGTGGCTGGCTGCTCAAGAGATTGATTGAAATGACTGTGTGAGTACGCTGCACACAGGA 1054
 Db 749 TAAGTGTAGGTGCTGCTACGCGTATGCTGGAAGATTCCTACAATATGCCGTTGAGCGCA 808
 QY 1055 AACAGTTTAACAGAGGCTCAGTGAATTTGGATTGATTGATTAGATGCTGCCGACTTGGTGA 1114
 Db 809 AACAGTTTGGTCAAGCGATTCGCAACTTCCAGTTGATTCAAGGTATGTTAGCCGATTTCTA 868
 QY 1115 CTCAGAAGGCTTACGCTCATGGAGAGTATGACCTACCTCACAGAGGATGTGTGAACCAAC 1174
 Db 869 AAGCTGAAATTTACGACGCAAAATGATGTTAGATGCTGCCGACTTGGTGA----- 923
 QY 1175 CTGGCTTTCCGACTGCTCCATCGAGGAGCCATCGTGAAGTGTTCAGCTCCGAGGCG 1234
 Db 924 -TGCTGGACAGAAATGTCAAGCGAAGCATCTTGTGCCAAGATGTTGCCACTGAAATGT 982
 QY 1235 CTTGGCAGTGTGTGAGTGGCGCTGCAGATCCTCGGGGGCTTGGGCTACACAGGAGT 1294
 Db 983 GTGGCGGTGTCGACAGATCGTGGCTACAGATCCATGTTGTCGGGTTATATCATGAT 1042
 QY 1295 ATCCGTACGAGCGCATACTGCTGCAGACCCGCGATCCTCTCATCTTCGAGGGAACCAATG 1354
 Db 1043 ATGCTATTGAGCGTTTTACCGTGTATGATGCTTTATTCGTTTGTATGAAGGTGACACGC 1102
 QY 1355 AGATTCTCGGATGATACATCGCC 1378
 Db 1103 AAATCCACAGGTCATATTGCCCC 1126

RESULT 14

US-09-648-004-27
 ; Sequence 27, Application US/09648004
 ; Patent No. 6498242
 ; GENERAL INFORMATION:
 ; APPLICANT: CHEN, QIONG
 ; APPLICANT: THOMAS, STUART
 ; APPLICANT: NAGARAJAN, VASANTHA
 ; TITLE OF INVENTION: BIOLOGICAL METHOD FOR THE PRODUCTION OF ADIPIC ACID AND
 ; TITLE OF INVENTION: INTERMEDIATES
 ; FILE REFERENCE: CL-1341-A
 ; CURRENT APPLICATION NUMBER: US/09/648,004
 ; CURRENT FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/252,553
 ; PRIOR FILING DATE: 1999-02-19
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 27
 ; LENGTH: 17417
 ; TYPE: DNA
 ; ORGANISM: Acinetobacter sp.
 US-09-648-004-27

Query Match 4.5%; Score 111.2; DB 4; Length 17417;
 Best Local Similarity 47.8%; Pred. No. 1.6e-20;
 Matches 499; Conservative 0; Mismatches 518; Indels 27; Gaps 5;
 QY 341 AAATCCAGATGAACCTTTGGAGAAATTAAGAGACCTTAGGCGCTTTTGGGCTGCAAGTCC 400

Db	1880	AAATTCAGCTGAAATCGTGACAGAAATGAAAGAACTGGGTCTTTTGGCTCACCAATTC	1939
QY	401	CAGAAGATATGCTGGCCTTGGCTTCTCCA--ACACCATCTACTCAGACTAGGGGAGA	457
Db	1940	CTGAGGAATAGAGGCTCTTGGCCTTGACCATGGAGGAAGAGTTTACATTGCAATTGAAC	1999
QY	458	TCATCAGCATGATGGGTCCATCACTGTGACCCCTGGCAGCGCACAGGCTATTGGCCCTCA	517
Db	2000	TGGAGCTACCTCTCCTCTTCCGTTCTCACTGATCGGCACCTAACATGGGATCGGTTTCA	2059
QY	518	AGGGGATCATCTTGGCTGGCACTGAGGAGCAGAAAGCCAAATACTTGCCTTAACTGCGGT	577
Db	2060	CAGGCTTAATTATTGATGGCTCCGAGACAGAACAGTATTTTTCGCCACGCTCTGGCAA	2119
QY	578	CCGGGAGCACATTCGACCCCTTTCGCTCACGGAGCCAGCCAGTGGGAGCGATGCGACGCT	637
Db	2120	GTGGTGAATATTGGTTTCATCTCTGTTTAACTGAACCTGATTCGGGTTTCAGATGCTGCCT	2179
QY	638	CAATCCGGAGGAGGCCACACTAAGTGAAGACAGAAGCACTACATCCTCAATGGGTCCA	697
Db	2180	CTTT-----AAAACCACACGGGTGAAGATGGTGATCATACATTTTAAATGGCACTA	2233
QY	698	AGGCTCGATTACTAATGGAGGACTGCCAATATTTTACTGTGTTTGCAGGACTGAGG	757
Db	2234	ACGCTTACATCACCATAACGATGCACCGATCGGGGTGCTTTACTGTCATGGCACGTAC	2288
QY	758	TCGTTGATTCCTGATGGATCAGTGAAGACACAAATCACAGCATTCATAGTAGAAGAGACT	817
Db	2289	----CAGTACCGAATTTAAGGTTACAGGTGGAATTTTCAGCCTTTATCTGTCAGGTAA	2344
QY	818	TTGGTGGAGTCACCTAATGGGAAACCCGAAGATAAATTAGGCATTCGGGGCTCCACACAT	877
Db	2345	CTCCTGGTATTTCTTGGGTAAACGCTGATAAGAAGATGGSCCAAAAAGGTGCACATAC	2404
QY	878	GTGAAGTCCATTTTGAACACCAACAGATACCTG---TGGAACATCTCTTGGAGAGGTG	934
Db	2405	GTGATGTGATTTTGAACATGTGCTATTCCTGCATCTGCATCTATGCTGTTGTTGAAG	2464
QY	935	GAGATGGGTTTAAGGTGGCCATGAACATCCTCAACAGCGCGCGTTTCAGCATGGCAGCG	994
Db	2465	GTGTAGGTTTAAACCTGCAATGAAGTACTTGTATAAGGCGGTATTCATATGCTGCAT	2524
QY	995	TCGTGGCTGGCTGCTCAAGAGATTGATTTGAATGACTCTGAGTACGCGCTGCACAGGA	1054
Db	2525	TAAGTGTAGTGTGCTACGCGTATGCTGGAAGATTCCCTACAATATGCGTTGAGCGCA	2584
QY	1055	AACAGTTTAAAGAGGCTCAGTGAATTTGGATTGATTCAGAGAGAAATTTGCACGTATGG	1114
Db	2585	AACAGTTTGGTCAAGCGATTGCGAACTCCAGTTGATTCAGGATGTTTATGCCGATTTCTA	2644
QY	1115	CTCAGAAGGCTTACGTCATGGAGAGTATGACCTACTCACAGCGGGATGTGGACCAAC	1174
Db	2645	AACTGNAATTTACGCAGCAAAATGTATGGTATTAGATGCTGCCGACTTCTGTGA-----	2699
QY	1175	CTGGCTTCCGACTCTCCATCGAGGACCGCATGGTGAAGGTGTTACGCTCCGAGCCG	1234
Db	2700	-TGCTGGACAGATGTCAGCAGGAAGCATCTGTGGCAAGATGTTTCCCACTGAAATGT	2758
QY	1235	CCGTGGCAGTGTGAGTGAGGCGCTGCAATCCTCGGGGGTGTGGCTTACACAGGACT	1294
Db	2759	GTGCGCGTGTGCGAGATCGTGGCTACAGATCCATGTGTGGTGGCGGTATATCAGTGAAT	2818
QY	1295	ATCCGTACGAGCGCATCTGCGCTGACACCCGATCCTCCTCATCTCTCGAGGGAACCAATG	1354
Db	2819	ATGCTATTGAGCGTTTTTACCCTGATGTACGTTTATCCGTTGTATGAAGGTACAACGC	2878
QY	1355	AGATTCTCCGGATGTACATCGCCC	1378
Db	2879	AAATCAACAGGTCATTTATGCCC	2902

US-09-328-352-2254
 ; Sequence 2254, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GYC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 2254
 ; LENGTH: 1170
 ; TYPE: DNA
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-2254

	Query Match	4.3%	Score 105.8;	DB 4;	Length 1170;
	Best Local Similarity	47.3%;	pred. No. 1.2e-19;		
	Matches 502;	Conservative 0;	Mismatches 532;	Indels 27;	Gaps 5;
QY	323	AAATTGACCAAGGAGGAATAATCCACATGAAACTTTGGAGAAAATTTGAAGAGCCTTAGGC	382		
Dd	101	AACGTGTAGAAGAAGACAAAATTCACAGATGACATTTGCCAAGAAATCGTGAGCTGGT	160		
QY	383	TTTTTGGCTGCAAGTCCCAGAGAATAATGTTGGCTTGGCTTCTCCAACACCATTGTACT	442		
Dd	161	TATTTGGCTACAATATCCCGAAGAAATATGTTGGCTTAGCATCACATATGGAAGAGGAAG	220		
QY	443	CRAAGACT--AGGGAGATCATCAGCATGGATGGTCCATCATCTGACCCCTGGCAGCGC	499		
Dd	221	TCAGAGTTGCATTTGAATTTGGACATTTGGACAACTTCACGGCTTTTTCGTCTCTTAATTGGTACCA	280		
QY	500	ACCAGCTATTGGCCCTCAAGGGGATCATCTTTGGCTGGCAGTGGAGGACAGAACCCCAAT	559		
Dd	281	ATAACGGTATTTGGCTCAAGTGCATTTTGAATGATGTCACAGAGACGAAAGCAGAAAT	340		
QY	560	ACTTTGCCTAAACTGGCGTGGGGAGGCATTTGCAAGCCCTTCTGCCTCACGGAGCAGCCA	619		
Dd	341	ATTTACCAAGTTACGCAAGTGGCGAAATATTTGGTTCAATTTGTTTAAACGAGCTGAAT	400		
QY	620	GTGGGAGCCATGCAGGCTCAATCCGAGCAGAGCCACACATTAAGTGAAGACAGNAGCACT	679		
Dd	401	CTGGTTCATGCTGCCTCTTTAAAAACAGTCCCGTA-----AAAGATGGCGATTTTT	454		
QY	680	ACATCCTCAATGGCTCCCAAGGCTGTGATTTACTAATGGAGACTGGCCAATATTTTACTG	739		
Dd	455	ATGTATTAAACGGCACCAAACGTTTATTACCAATGCACCCCATGCTGCAACTTTTACCG	514		
QY	740	TGTTTTGCAAGACTGAGGTCGTTGATCTCATGTGGATCAGTGAAGACAAANATCACAGCAT	799		
Dd	515	TAATGGCTCGGAC-----GAATCTCTGAAATTAAGGGGTGGGTGGCATTTTCAGCTT	565		
QY	800	TCATAGTAGAAGAGACTTTGTTGGAGTCACCTAATGGGAACCCGAAGATTAATATAGGCA	859		
Dd	566	TTTTAGTCGAGGCCAATACACCGGCATTTACATTAGGCAAAATTGACCAAAATGGGGC	625		
QY	860	TTGCGGGCTCCAACACTTGTGAAGTCCATTTTGAANAACACCAAGATACCTG---TGAA	916		
Dd	626	AAAAGGGTTACATACCTGTGATGTCAATTTTTTGAANAATTTGCGGTGTACCTTCTGCAC	685		
QY	917	ACATCTTTGGAGAGGTGCGAGATGGGTTTAAGTGGCCATGAACATCCCTCAACAGCGGC	976		
Dd	686	TGATTTGGTGGCGTAGAAGGTGTCGGTTTAAACGGCCATGAAGTCTGGATTAAGGGC	745		
QY	977	GGTTACGATGGGAGCGTCTGGCTGGGCTGCTCAAGAGATTGATTAAGATGACTGCTG	1036		
Dd	746	GTTTTCGACATTTGGTCCCTATAGCTGGGTGTTGCCGAACGCCATGTTAAACGACGCATTC	805		
QY	1037	AGTACGCTGACACAGGAACACAGTTTAAACAGAGGCTCAGTCAATTTGGATTGATTCAGG	1096		
Dd	806	ACTATGCGATTGACCGTAGCAGTTTGGTCAAGCCTATTTCGAACTTTTACGCTTATTCAG	865		

Tue Sep 9 10:14:36 2003

```
QY 1097 AGAATTTGCACTGATGGCTCAGAAGGCTTAGCTCATGGAGAGTATGACCTACCTCACAG 1156
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 866 CCATGCTTCCCGACTCAAAAGCCGAGATTATGACAGCTAAATGTAAGTCTCTAGATGCAG 925
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1157 CAGGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGCGACCCATGGTGAAGG 1216
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 926 CTCGCCCGCTGATATGCG- ----AAACATTAGTACAGAAGCATCTTGGCGCAAAA 979
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1217 TGTTCAGCTCCGAGGCCCTGGCAGTGTGTGAGTGAGGCGCTGCAGATCCTCGGGGCT 1276
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 980 TGTTCGCACTGAAATGTGGGACGTGTTGCTGACCGCTGTGTACAGATTCACGGTGGTG 1039
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1277 TGGGCTACACAGGGGACTATCCGTACGAGCGCATACTGCGTGCACACCGCATCCTCCTCA 1336
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1040 CGGGCTATATCAGTGAATGCGATGACCGTTTTATCGAGATGCGGTTGTTCCGTT 1099
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1337 TCTTCGAGGGAACCAATGAGATTCCTCCGATGTACATCGCC 1377
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1100 TATATGAAGGAACCAACGCAAGTTCAACAACATTATTATGCC 1140
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: September 6, 2003, 12:58:24
Job time : 154.926 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2003, 23:17:14 : Search time 554.612 Seconds
(without alignments)
10168.445 Million cell updates/sec

Title: US-09-945-326-1
Perfect score: 2452
Sequence: 1 cgtgtgtgtcctcgcgc.....actgttaaaaaaaaaaaaaa 2452

Scoring table: IDENTITY_NUC

Searched: 1537136 seqs, 1149988732 residues

Total number of hits satisfying chosen parameters: 3074272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2452	100.0	2452	10	US-09-945-326-1
2	2438.4	99.4	2440	14	US-10-168-274-51
3	1863	76.0	1863	10	US-09-945-326-3
4	1102.4	45.0	2080	14	US-10-198-846-11756
5	457.6	18.7	502	11	US-09-918-995-20366
6	452.2	18.4	490	11	US-09-918-995-21901
7	451.8	18.4	483	11	US-09-918-995-23707
8	390.6	15.9	451	11	US-09-917-800A-11034
9	388.6	15.8	2117	10	US-09-918-995-1584
10	261.4	10.7	490	11	US-09-918-995-16927
11	258.6	10.5	312	9	US-09-822-849A-480
12	187.4	7.6	402	11	US-09-918-995-4541
13	165.2	6.7	568	10	US-09-796-692-8175
14	165.2	6.7	568	14	US-10-040-862-8175
15	159.4	6.5	183	10	US-09-867-701-3392
16	151.2	6.2	1158	14	US-10-156-761-5259

c	17	151.2	6.2	9025608	14	US-10-156-761-1	Sequence 1, Appli
	18	147.6	6.0	1170	14	US-10-156-761-6897	Sequence 6897, Ap
	19	147.6	6.0	9025608	14	US-10-156-761-1	Sequence 1, Appli
	20	146.6	6.0	1829	10	US-09-880-107-2309	Sequence 2309, Ap
	21	132.6	5.4	1155	14	US-10-156-761-5010	Sequence 5010, Ap
	22	126.6	5.2	1716	10	US-09-974-300-1101	Sequence 1101, Ap
	23	120.8	4.9	1138	10	US-09-974-300-1082	Sequence 1082, Ap
	24	119.8	4.9	1149	14	US-10-156-761-6577	Sequence 6577, Ap
	25	118.2	4.8	1866	10	US-09-917-800A-1617	Sequence 1617, Ap
	26	115.6	4.7	810	10	US-09-974-300-5530	Sequence 5530, Ap
	27	114.4	4.7	1155	14	US-10-156-761-5264	Sequence 5264, Ap
	28	111.2	4.5	1155	14	US-10-272-419-5	Sequence 5, Appli
	29	111.2	4.5	17417	14	US-10-272-419-27	Sequence 27, Appli
	30	110.6	4.5	1056	10	US-09-974-300-5546	Sequence 5546, Ap
	31	105	4.3	1938	14	US-10-156-761-2554	Sequence 2554, Ap
c	32	104.2	4.2	411	10	US-09-960-352-1781	Sequence 1781, Ap
	33	103.6	4.2	439	10	US-09-998-598-317	Sequence 317, App
	34	102.4	4.2	495	11	US-09-918-995-350	Sequence 350, App
	35	101.8	4.2	439	10	US-09-998-598-317	Sequence 171, Ap
	36	100.2	4.1	562	10	US-10-060-036-766	Sequence 766, App
	37	94.6	3.9	2238	14	US-10-080-823-3	Sequence 1171, Ap
	38	92.6	3.8	983	10	US-09-974-300-1090	Sequence 3, Appli
	39	90	3.7	2682	10	US-09-880-107-3300	Sequence 1090, Ap
	40	90	3.7	2682	12	US-10-354-358-27	Sequence 3300, Ap
	41	86.8	3.5	466	10	US-09-764-847-110	Sequence 27, Appli
	42	86.8	3.5	466	14	US-10-092-154-110	Sequence 110, App
c	43	84.6	3.5	650	13	US-10-027-632-229219	Sequence 110, App
	44	81.8	3.3	615	10	US-09-974-300-1102	Sequence 229219,
	45	81.4	3.3	1383	10	US-09-925-300-705	Sequence 1102, Ap
							Sequence 705, App

ALIGNMENTS

RESULT 1
US-09-945-326-1
; Sequence 1, Application US/09945326
; Patent No. US20020127680A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Hunter, John Joseph
; TITLE OF INVENTION: 62112, A NOVEL HUMAN DEHYDROGENASE AND
; FILE REFERENCE: MNI-187
; FILE REFERENCE: US/09/945,326
; CURRENT APPLICATION NUMBER: 2001-08-31
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/229,831
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)...(1932)
US-09-945-326-1

Query Match	100.0%	Score 2452;	DB 10;	Length 2452;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2452;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	CGTGTGTGTGTCCTCGCGCGCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGAACATCG	60	
DB	1	CGTGTGTGTGTCCTCGCGCGCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGAACATCG	60	
QY	61	GGCAGCATGAGCGGCTGCGGCTCTTCTGCGCACCGGCTGCGGCTGCGCTGCGCGG	120	
DB	61	GGCAGCATGAGCGGCTGCGGCTCTTCTGCGCACCGGCTGCGGCTGCGCTGCGCGG	120	
QY	121	GSTCTGTGGTCTCTACCGCGAACCGGGGCTACTGCGCACCGGCTGCTACGAGCT	180	

Db 121 GGTCTGTGGTCTCTACCGCGAACCGCGGGCTACTCGGCACACCGCCGCTCTACGAGCT 180
Qy 181 TTCCCCAAAGAGCTTTTCTCTAGGCAAAATCAAGAGAAAGAAAGTTTCCCATTTCCAGAA 240
Db 181 TTCCCCAAAGAGCTTTTCTCTAGGCAAAATCAAGAGAAAGAAAGTTTCCCATTTCCAGAA 240
Qy 241 GTTAGCCAAAGATGAACATTAATGAATCAATCAGTCTTGGGACCGCTGGAAAAATCTTTC 300
Db 241 GTTAGCCAAAGATGAACATTAATGAATCAATCAGTCTTGGGACCGCTGGAAAAATCTTTC 300
Qy 301 ACTGAAGAGTGGACTCCCAAAATTTGACAGGAAGGAAATCCAGATGAACATTTTG 360
Db 301 ACTGAAGAGTGGACTCCCAAAATTTGACAGGAAGGAAATCCAGATGAACATTTTG 360
Qy 361 GAGAAATTTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCCAAGAAATATGTTGGCCCTG 420
Db 361 GAGAAATTTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCCAAGAAATATGTTGGCCCTG 420
Qy 421 GGCCTTCCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGTCCATC 480
Db 421 GGCCTTCCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGTCCATC 480
Qy 481 ACTGTGACCCCTGGCAGCGCACAGGCTATTGGGCTCAAGGGGATCATCTTGGCTGGCACT 540
Db 481 ACTGTGACCCCTGGCAGCGCACAGGCTATTGGGCTCAAGGGGATCATCTTGGCTGGCACT 540
Qy 541 GAGGAGCAGAAAGCCAAATCTGCTAAACTGSCGTCGCGGAGCACAATTCGAGCCCTC 600
Db 541 GAGGAGCAGAAAGCCAAATCTGCTAAACTGSCGTCGCGGAGCACAATTCGAGCCCTC 600
Qy 601 TGCCCTCAGGAGCCAGCCAGTGGGAGCGATGACGCCCTCAATTCGGAGCAGAGCCACTA 660
Db 601 TGCCCTCAGGAGCCAGCCAGTGGGAGCGATGACGCCCTCAATTCGGAGCAGAGCCACTA 660
Qy 661 AGTGAAGACAGAAGCACTACATCTCAATGCTCCCAAGGTCGAGTCTGATTAATGGAGGA 720
Db 661 AGTGAAGACAGAAGCACTACATCTCAATGCTCCCAAGGTCGAGTCTGATTAATGGAGGA 720
Qy 721 CTGGCCAAATATTTTACTGTGTTTGCAAAGACTGAGGTCGTTGATCTGTGATGATCAGTG 780
Db 721 CTGGCCAAATATTTTACTGTGTTTGCAAAGACTGAGGTCGTTGATCTGTGATGATCAGTG 780
Qy 781 AAAGACAAATCACAGCATTCATAGTAGAAGAGACTTTGTGGAGTCACCTAAATGGGAAA 840
Db 781 AAAGACAAATCACAGCATTCATAGTAGAAGAGACTTTGTGGAGTCACCTAAATGGGAAA 840
Qy 841 CCGGAAGATAAATTAGGCATTCCGGGCTCCAAACACTTGTGAAGTCCATTTTGAACACC 900
Db 841 CCGGAAGATAAATTAGGCATTCCGGGCTCCAAACACTTGTGAAGTCCATTTTGAACACC 900
Qy 901 AAGATACCTGTGGAACATCCTTGGAGAGTCCGAGATGGGTTTAAAGTGGCCATGAAC 960
Db 901 AAGATACCTGTGGAACATCCTTGGAGAGTCCGAGATGGGTTTAAAGTGGCCATGAAC 960
Qy 961 ATCCTCAACAGCGCCGGTTACAGATGGCAGCGCTGCTGGCTGGCTGCTCAAGAGATTG 1020
Db 961 ATCCTCAACAGCGCCGGTTACAGATGGCAGCGCTGCTGGCTGGCTGCTCAAGAGATTG 1020
Qy 1021 ATTGAATGACTGCTGAGTACCGCTGCACAAAGAAACAGTTTAAAGAGGCTCAGTGAA 1080
Db 1021 ATTGAATGACTGCTGAGTACCGCTGCACAAAGAAACAGTTTAAAGAGGCTCAGTGAA 1080
Qy 1081 TTTGGATTGATTTCAGGAGAAATTTGCACATGATGGCTCAGAAGGCTTACGTCATGGAGAT 1140
Db 1081 TTTGGATTGATTTCAGGAGAAATTTGCACATGATGGCTCAGAAGGCTTACGTCATGGAGAT 1140
Qy 1141 ATGACCTACCTCACAGAGGATGCTGGACCAACCTGGCTTTCCGACTGCTCCATCGAG 1200
Db 1141 ATGACCTACCTCACAGAGGATGCTGGACCAACCTGGCTTTCCGACTGCTCCATCGAG 1200
Qy 1201 GCAGCCATGTTGAAGGTTTACGCTCCGAGCGCCGCTGGCAGTGTGTGAGTGAAGGCGCTG 1260
Db 1201 GCAGCCATGTTGAAGGTTTACGCTCCGAGCGCCGCTGGCAGTGTGTGAGTGAAGGCGCTG 1260

Qy 1261 CAGATCCTCGGGGCTTGGGCTACACAAGGAGACTATCCGTACGAGGCGCATACTCTGCGTAC 1320
Db 1261 CAGATCCTCGGGGCTTGGGCTACACAAGGAGACTATCCGTACGAGGCGCATACTCTGCGTAC 1320
Qy 1321 ACCCGCATCTCTCATCTCTCGAGGGAACCAATGAGATTCTCGGATGTACATCGCCCTG 1380
Db 1321 ACCCGCATCTCTCATCTCTCGAGGGAACCAATGAGATTCTCTCGGATGTACATCGCCCTG 1380
Qy 1381 ACGGTCTGACAGTACCGCGCGCATCTCTGACTACCAAGATCCATGAGCTTAAACAGGCC 1440
Db 1381 ACGGTCTGACAGTACCGCGCGCATCTCTGACTACCAAGATCCATGAGCTTAAACAGGCC 1440
Qy 1441 AAAGTGAACACAGTCATGATACCGTTGGCCGAGGCTTCGGGACTCCCTGGGCGGCAACT 1500
Db 1441 AAAGTGAACACAGTCATGATACCGTTGGCCGAGGCTTCGGGACTCCCTGGGCGGCAACT 1500
Qy 1501 GTGGACCTGGGCTGACAGGCAACCATGAGCTTGTGACCCCACTCTTTCGGGAGAGTGCC 1560
Db 1501 GTGGACCTGGGCTGACAGGCAACCATGAGCTTGTGACCCCACTCTTTCGGGAGAGTGCC 1560
Qy 1561 AAACAAGTTTGAAGGAACACCTTACTGTTCGGCGGAGCCGTTGGAGACACTGCTCTCCG 1620
Db 1561 AAACAAGTTTGAAGGAACACCTTACTGTTCGGCGGAGCCGTTGGAGACACTGCTCTCCG 1620
Qy 1621 TTTGSCAAGACCATCATGAGGAGCAGCTGTGACTGAAGCGGTTGGCCCAACATCCTCATC 1680
Db 1621 TTTGSCAAGACCATCATGAGGAGCAGCTGTGACTGAAGCGGTTGGCCCAACATCCTCATC 1680
Qy 1681 AACCTGTATGGCATGACGCCCTGTGTGCGGGGCGACCGCTCCATCCGATTGGGCTC 1740
Db 1681 AACCTGTATGGCATGACGCCCTGTGTGCGGGGCGACCGCTCCATCCGATTGGGCTC 1740
Qy 1741 CGCAACACGACCAAGAGTTCTTGTGGCCAAACACTTCTGCTGCGTGAAGCTTACTTTGCGAG 1800
Db 1741 CGCAACACGACCAAGAGTTCTTGTGGCCAAACACTTCTGCTGCGTGAAGCTTACTTTGCGAG 1800
Qy 1801 AATCTCTTCAGGCTCTCTCAGCTGGACAAAGTATGCTCCAGAAACCTTAGTAGAGCAGATT 1860
Db 1801 AATCTCTTCAGGCTCTCTCAGCTGGACAAAGTATGCTCCAGAAACCTTAGTAGAGCAGATT 1860
Qy 1861 AAGAAAGTGTCCAGCAGATCCTTGAAGAGCAGCCTATATCTGTGCCCACTCTGGAC 1920
Db 1861 AAGAAAGTGTCCAGCAGATCCTTGAAGAGCAGCCTATATCTGTGCCCACTCTGGAC 1920
Qy 1921 AGGACATGCTGAGGCGAGGAGACAGTGTCCCTGCTACCGCCGCGCCCTACCCATGGCCG 1980
Db 1921 AGGACATGCTGAGGCGAGGAGACAGTGTCCCTGCTACCGCCGCGCCCTACCCATGGCCG 1980
Qy 1981 TTGCTGGATGACTGTACTCTTTTTCAGAAAGTGTGGGATTTATCAGAGTTAAGCCTT 2040
Db 1981 TTGCTGGATGACTGTACTCTTTTTCAGAAAGTGTGGGATTTATCAGAGTTAAGCCTT 2040
Qy 2041 TTGTTCCCGCTGCACTGAAGGTTGTCCCTGGCCTGGGAGAGCCCTCTCCAGGTTT 2100
Db 2041 TTGTTCCCGCTGCACTGAAGGTTGTCCCTGGCCTGGGAGAGCCCTCTCTCCAGGTTT 2100
Qy 2101 TGACCTGCAGGAGTGTCTCTACAGGACCATCAGAGTCTTGAAGTGAAGCCGAGAGA 2160
Db 2101 TGACCTGCAGGAGTGTCTCTACAGGACCATCAGAGTCTTGAAGTGAAGCCGAGAGA 2160
Qy 2161 GAGAAATGGAATGCTGACCCCTGGAAGTGGCGGATTTCTGCTGATTTAGGAGACACCAT 2220
Db 2161 GAGAAATGGAATGCTGACCCCTGGAAGTGGCGGATTTCTGCTGATTTAGGAGACACCAT 2220
Qy 2221 AGTGAAGTGGGCTTATGCTGCTCCAGGTTGAGGTTGGGAGCTGTGTC 2280
Db 2221 AGTGAAGTGGGCTTATGCTGCTCCAGGTTGAGGTTGGGAGCTGTGTC 2280
Qy 2281 AGGTGGAATGAGCAATTTCTCTCAACACACATTTCTTAAGAAACAGCTTGAAGCTCT 2340
Db 2281 AGGTGGAATGAGCAATTTCTCTCAACACACATTTCTTAAGAAACAGCTTGAAGCTCT 2340

		CURRENT FILING DATE: 2001-08-31		PRIOR APPLICATION NUMBER: 60/229,831	
		PRIOR FILING DATE: 2000-08-31		NUMBER OF SEQ ID NOS: 3	
		SOFTWARE: FastSeq for Windows Version 4.0		SEQ ID NO 3	
		LENGTH: 1863		TYPE: DNA	
		ORGANISM: Homo sapiens		FEATURE:	
		NAME/KEY: CDS		LOCATION: (1)....(1863)	
		US-09-945-326-3			
		Query Match		76.0%; Score 1863; DB 10; Length 1863;	
		Best Local Similarity		100.0%; Pred. No. 0;	
		Matches 1863; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1568	TTGAGGAGAACACCTACTGCTTTGGCGGACCGTGGAGACACTGCTGCTCCGCTTTGGCA	1627	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	126
Db	1561	TTGAGGAGAACACCTACTGCTTTGGCGGACCGTGGAGACACTGCTGCTCCGCTTTGGCA	1620	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	60
QY	1628	AGACCATATGAGGAGAGCTGGTACTGAACGGGTGGCCAACTCTCATCAACCTGT	1687	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	186
Db	1621	AGACCATATGAGGAGAGCTGGTACTGAACGGGTGGCCAACTCTCATCAACCTGT	1680	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	120
QY	1688	ATGGCATGACGGCGCTGCTGCGGGCCAGCGCTCCATCCGCAATGGGCTCCGCAACC	1747	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	186
Db	1681	ATGGCATGACGGCGCTGCTGCGGGCCAGCGCTCCATCCGCAATGGGCTCCGCAACC	1740	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	120
QY	1748	ACGACCAAGAGGTTCTCTTTGGGCAACACCTTCTGGTGAAGCTTACTTGCAGATCTCT	1807	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	186
Db	1741	ACGACCAAGAGGTTCTCTTTGGGCAACACCTTCTGGTGAAGCTTACTTGCAGATCTCT	1800	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	120
QY	1808	TCAGCCTCTCTAGCTGGACAGTATGCTCCAGAAAACCTAGATGAGCAGATTAAGAAG	1867	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	186
Db	1801	TCAGCCTCTCTAGCTGGACAGTATGCTCCAGAAAACCTAGATGAGCAGATTAAGAAG	1860	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	120
QY	1868	TGTCGCCAGCAGATCCTTTGAGAGCGAGCTATATCTGTGCCACCTCTGGACAGACAT	1927	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	186
Db	1861	TGTCGCCAGCAGATCCTTTGAGAGCGAGCTATATCTGTGCCACCTCTGGACAGACAT	1920	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	120
QY	1928	GCTGAGCGAGGACAGTGTCCCTGCTACCGCCCGCCCTACCCATGCGCCGCTTGTGG	1987	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	186
Db	1921	GCTGAGCGAGGACAGTGTCCCTGCTACCGCCCGCCCTACCCATGCGCCGCTTGTGG	1980	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	120
QY	1988	ATGACTCTTACTCTTTTTCAGAGGTTGGGATATACAGAGTTAAGCTTTTGTTC	2047	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	186
Db	1981	ATGACTCTTACTCTTTTTCAGAGGTTGGGATATACAGAGTTAAGCTTTTGTTC	2040	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	120
QY	2048	CCGTCTGACCTGAAGGTTGTCGCTGCGCTGGGAGGCTCTTCCAGTTTGTGACCTG	2107	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	186
Db	2041	CCGTCTGACCTGAAGGTTGTCGCTGCGCTGGGAGGCTCTTCCAGTTTGTGACCTG	2100	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	120
QY	2108	CAGGAGTGCTCTCTAACAGAGCACCACAGCTTCTGAAGTGAAGCGGAGAGAGATG	2167	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	186
Db	2101	CAGGAGTGCTCTCTAACAGAGCACCACAGCTTCTGAAGTGAAGCGGAGAGAGATG	2160	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	120
QY	2168	GAATGCTGACCTGAACTGGCGGTATCTGCTATGAGGAGACACCATAGTGAA	2227	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	186
Db	2161	GAATGCTGACCTGAACTGGCGGTATCTGCTATGAGGAGACACCATAGTGAA	2220	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	120
QY	2228	ACTGGGCTTATGCTGCTCCAGGCTGTGAGGTGGGTGGGACCTGTGTCAAGTGTG	2287	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	186
Db	2221	ACTGGGCTTATGCTGCTCCAGGCTGTGAGGTGGGTGGGACCTGTGTCAAGTGTG	2280	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	120
QY	2288	GATAGCCATTTCTGCTCAACACACATCTCTAAGAACAGCTTGAAGCTCTGCTGGG	2347	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	186
Db	2281	GATAGCCATTTCTGCTCAACACACATCTCTAAGAACAGCTTGAAGCTCTGCTGGG	2340	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	120
QY	2348	TCATTCTTTAACTAGAGCAGAGGCACTTAAACATGTACAGGAGACCATTTAAACAA	2407	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	186
Db	2341	TCATTCTTTAACTAGAGCAGAGGCACTTAAACATGTACAGGAGACCATTTAAACAA	2400	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	120
QY	2408	GAATATAAATGTCACATCTGTGTACTGTTTAAAAA	2447	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	186
Db	2401	GAATATAAATGTCACATCTGTGTACTGTTTAAAAA	2440	ATGAGCGGCTCGGGGCTCTTCTCGGCACACGCTGCGGCTCGCTGCTGCCGGGCTG	120
		RESULT 3		US-09-945-326-3	
		Sequence 3, Application US/09945326		Patent No. US20020127680A1	
		GENERAL INFORMATION:		APPLICANT: Meyers, Rachel	
		APPLICANT: Hunter, John Joseph		TITLE OF INVENTION: 62112, A NOVEL HUMAN DEHYDROGENASE AND	
		TITLE OF INVENTION: USES THEREOF		FILE REFERENCE: MNI-187	
		CURRENT APPLICATION NUMBER: US/09/945,326			


```
QY 907 CCTGTGGAACATCCTTTGGAGAGCTCGGAGATGGGTTTAAGTTGGCCATGAACATCCTC 966
Db 841 CCTGTGGAACATCCTTTGGAGAGCTCGGAGATGGGTTTAAGTTGGCCATGAACATCCTC 900
QY 967 AACAGCGCGGTTTACAGATGGGAGCGTCTGGCTGGCTCAAGAGATTCATTGAA 1026
Db 901 AACAGCGCGGTTTACAGATGGGAGCGTCTGGCTGGCTCAAGAGATTCATTGAA 960
QY 1027 ATGACTGCTGAGTACGCTGCACAAAGAAACAGTTTAAACAGAGCTCAGTGAATTGGA 1086
Db 961 ATGACTGCTGAGTACGCTGCACAAAGAAACAGTTTAAACAGAGCTCAGTGAATTGGA 1020
QY 1087 TTGATTACAGGAAATTTGCACTGATGGCTCAGAGGCTTACGTCATGGAGATGATGACC 1146
Db 1021 TTGATTACAGGAAATTTGCACTGATGGCTCAGAGGCTTACGTCATGGAGATGATGACC 1080
QY 1147 TACCTCACAGCAGGATGCTGGACCAACCTTGGCTTTCCGACTGCTCCATCGAGGAGCC 1206
Db 1081 TACCTCACAGCAGGATGCTGGACCAACCTTGGCTTTCCGACTGCTCCATCGAGGAGCC 1140
QY 1207 ATGCTGAAGTGTTCAGCTCCGAGCGGCTGCGAGTGTGAGTGAAGGCTGCGAGATC 1266
Db 1141 ATGCTGAAGTGTTCAGCTCCGAGCGGCTGCGAGTGTGAGTGAAGGCTGCGAGATC 1200
QY 1267 CTCGGGGCTTGGCTACAAAGGACTATCCGTACGAGCGCATCTTGGGTGACACCCCG 1326
Db 1201 CTCGGGGCTTGGCTACAAAGGACTATCCGTACGAGCGCATCTTGGGTGACACCCCG 1260
QY 1327 ATCTCTCTCATCTTCAGAGGAACCAATGAGATTTCCGGATGTACATCGCCCTGACGGT 1386
Db 1261 ATCTCTCTCATCTTCAGAGGAACCAATGAGATTTCCGGATGTACATCGCCCTGACGGT 1320
QY 1387 CTCAGCATCCCGCGCATCTCTGACTACAGGATCCATGAGCTTAAACAGGCCAAAGTG 1446
Db 1321 CTCAGCATCCCGCGCATCTCTGACTACAGGATCCATGAGCTTAAACAGGCCAAAGTG 1380
QY 1447 AGCAGAGTCATGATACCTTGGCCGAGGCTTCGGGACTCCCTGGCCGCAACTGTGGAC 1506
Db 1381 AGCAGAGTCATGATACCTTGGCCGAGGCTTCGGGACTCCCTGGCCGCAACTGTGGAC 1440
QY 1507 CTGGGCTGACAGGCAACCATGGAGTTGTGCACCCCGAGTCTTGGGACAGTCCCAACAG 1566
Db 1441 CTGGGCTGACAGGCAACCATGGAGTTGTGCACCCCGAGTCTTGGGACAGTCCCAACAG 1500
QY 1567 TTTGAGGAGAACACCTACTGCTTGGCCGAGCGTGGAGACTCTTGGGACAGTCCCAACAG 1626
Db 1501 TTTGAGGAGAACACCTACTGCTTGGCCGAGCGTGGAGACTCTTGGGACAGTCCCAACAG 1560
QY 1627 AAGACCATCATGGAGGAGCAGCTGTTACTGAAGCGGTTGGCCAGCATCTCTCATCAACCTG 1686
Db 1561 AAGACCATCATGGAGGAGCAGCTGTTACTGAAGCGGTTGGCCAGCATCTCTCATCAACCTG 1620
QY 1687 TATGGCATGACGGCGTGCTGCGGGCCAGCCCTCCATCCGATTTGGCTCCGCAAC 1746
Db 1621 TATGGCATGACGGCGTGCTGCGGGCCAGCCCTCCATCCGATTTGGCTCCGCAAC 1680
QY 1747 CACGACACAGAGGTTCTCTTGGCCAAACCTTCTCGGTGGAGCTTACTTCGAAATCTC 1806
Db 1681 CACGACACAGAGGTTCTCTTGGCCAAACCTTCTCGGTGGAGCTTACTTCGAAATCTC 1740
QY 1807 TTAGCCCTCTCTAGCTGGACAGTATGCTCCAGAAACCTTAGATGAGCAGATTAAAGAA 1866
Db 1741 TTAGCCCTCTCTAGCTGGACAGTATGCTCCAGAAACCTTAGATGAGCAGATTAAAGAA 1800
QY 1867 GTGTCCACAGCAGATCTTTGAGAGCGGCTTATATCTGTGCCCCACCCTCTGGACAGGACA 1926
Db 1801 GTGTCCACAGCAGATCTTTGAGAGCGGCTTATATCTGTGCCCCACCCTCTGGACAGGACA 1860
QY 1927 TGC 1929
Db 1861 TGC 1863
```

RESULT 4

US-10-198-846-11756
: Sequence 11756, Application US/10198846
: Publication No. US20030099974A1
: GENERAL INFORMATION:

: APPLICANT: Lillie, James
: APPLICANT: Xu, Yongyao
: APPLICANT: Wang, Youzhen
: APPLICANT: Steinmann, Kathleen
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
: TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
: TITLE OF INVENTION: THERAPY OF BREAST CANCER
: FILE REFERENCE: MRI-049
: CURRENT APPLICATION NUMBER: US/10/198,846
: PRIOR FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/306,220
: NUMBER OF SEQ ID NOS: 14084
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 11756
: LENGTH: 2080
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 2075, 2076, 2077, 2078, 2079, 2080
: OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11756

Query Match

Best Local Similarity 45.0%; Score 1102.4; DB 14; Length 2080;

Matches 1359; Conservative 0; Mismatches 81; Indels 38; Gaps 17;

```
QY 26 GAAGGGGAGACTGAGGCTGAGGCTGGGGAACATCGGCAGCATGAGCGCTGGGGCTCT 85
Db 450 GAAGGCCAGACTGAGGCTGAGGCTGGGGAACATCGGCAGCATGAGCGCTGGGGCTCT 509
QY 86 TCCTTGGCCACCACCGCTGGCGCTGCTGCTGCCGGGTCTGGTGTCTTACCGCAACC 145
Db 510 TCCTTGGCCACCACCGCTGGCGCTGCTGCTGCCGGGTCTGGTGTCTTACCGCAACC 569
QY 146 GCGGCTACTTGGCCACCAGCCGCTCTACGAGCTTTCGCCAAGAGCTTTTCTTCCAGGCA 205
Db 570 GCGGCTACTTGGCCACCAGCCGCTCTACGAGCTTTCGCCAAGAGCTTTTCTTCCAGGCA 639
QY 206 AATCAAGAGAAAGAGTTTCCCAATTTCCAGAAGTTAGCCAAAGATGAACATTAATGAA 265
Db 630 CAATCCCGAAGAG-----GTTAAGTTAGCCCAAGATGAACATTAATGAA 674
QY 266 TCAATCAGTTCTTGGGACCGCTGGAATAATTTCTTCACTGAAGGTGGACTCCCGAATA 325
Db 675 TCAATCAGTTCTTGGGACCGCTGGAATAATTTCTTCACTGAAGGTGGACTCCCGAATA 734
QY 326 TTGACCAAGAA - GGGAAATCCAGATGAACATTTTGGAGAA - TTCAAGAGCTAGGGCT 383
Db 735 TTGACCAAGAA - GGGAAATCCAGATGAACATTTTGGAGAA - TTCAAGAGCTAGGGCT 794
QY 384 TTTTGGGCTCC - AAGTCCCAAGAGATATGTTGGCTGGCTGGGCTTCT - CCAACACCATGTAC 441
Db 795 TTTTGGGCTCC - AAGTCCCAAGAGATATGTTGGCTGGCTGGGCTTCT - CCAACACCATGTAC 854
QY 442 TCAAGACTAGGGAGATCATCAGATGGATGGTTCATCAGTGTGAGTGGCTGGCGGCTGAC 501
Db 855 TCAAGACTAGGGAGATCATCAGATGGATGGTTCATCAGTGTGAGTGGCTGGCGGCTGAC 914
QY 502 CAGGCTA - TTGGCTCAAGGGGATCATCTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 560
Db 915 CAGGCTA - TTGGCTCAAGGGGATCATCTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 974
QY 561 CTTGCTTAACTGGGCTCGGGGAGCAGATTCAGGCTTCTGCTTACGAGGAGCCAGCCAG 620
Db 975 CTTGCTTAACTGGGCTCGGGGAGCAGATTCAGGCTTCTGCTTACGAGGAGCCAGCCAG 1034
```

```

RESULT 6
US-09-918-995-21901
? Sequence 21901, Application US/09918995
? Publication No. US20030073623A1
? GENERAL INFORMATION:
? APPLICANT: Hyseq, Inc.
? TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
? TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
? FILE REFERENCE: 20411-756
? CURRENT APPLICATION NUMBER: US/09/918,995
? CURRENT FILING DATE: 2001-07-30
? PRIOR APPLICATION NUMBER: US/09/235,076
? PRIOR FILING DATE: 1999-01-20
? NUMBER OF SEQ ID NOS: 38054
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 21901
? LENGTH: 490
? TYPE: DNA
? ORGANISM: Homo sapiens
US-09-918-995-21901
Query Match 18.4%; Score 452.2; DB 11; Length 490;

```

```

RESULT 5
US-09-918-995-20366
; Sequence 20366, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc..
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30

```

Best Local Similarity 99.3%; Pred. No. 4e-130;
Matches 454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 1415 CCAGATCCATGAGCTTAACAGGCCAAAGTGAGCACAGTCATGGATACCGTTGGCCGGA 1474
Db 33 CCAAGATCCATGAGCTTAACAGGCCAAAGTGAGCACAGTCATGGATACCGTTGGCCGGA 92
QY 1475 GGCCTTCGGGACTCCCTGGGCGGAACCTGTGGACCTGGGCTGACAGGCAACCATGGAGTTG 1534
Db 93 GGCCTTCGGGACTCCCTGGGCGGAACCTGTGGACCTGGGCTGACAGGCAACCATGGAGTTG 152
QY 1535 TGCACCCAGCTTTCGGGACAGTGCCCAACAAGTTTGAGGAGAACACCTACTGCTTCGGCC 1594
Db 153 TGCACCCAGCTTTCGGGACAGTGCCCAACAAGTTTGAGGAGAACACCTACTGCTTCGGCC 212
QY 1595 GGACCGTGGAGACACTGCTGCTCGCTTTCGCAAGACCATCATGGAGGAGCAGCTGGTAC 1654
Db 213 GGACCGTGGAGACACTGCTGCTCGCTTTCGCAAGACCATCATGGAGGAGCAGCTGGTAC 272
QY 1655 TGAAGCGGTGGCAACATCCTCATCAACCTGTATGGCATGACGGCGCTGCTGCGCGG 1714
Db 273 TGAAGCGGTGGCAACATCCTCATCAACCTGTATGGCATGACGGCGCTGCTGCGCGG 332
QY 1715 CCAGCCGCTCCATCCGATTGGGCTCGCAACACACAGGAGTTCTTGGCCCAACA 1774
Db 333 CCAGCCGCTCCATCCGATTGGGCTCGCAACACACAGGAGTTCTTGGCCCAACA 392
QY 1775 CCTTCTCGGTGGAGCTTACTTGCAGAACTCTTTCAGCCTCTCTCAGCTGGACAAGTATG 1834
Db 393 CCTTCTCGGTGGAGCTTACTTGCAGAACTCTTTCAGCCTCTCTCAGCTGGACAAGTATG 452
QY 1835 CTCAGAAAACCTTAGATGACGAGATTAAAGAAAGTGC 1871
Db 453 CTCAGAAAACCTTAGATGACGAGATTAAAGAAAGTGC 489
```

RESULT 7

US-09-918-995-23707
; Sequence 23707, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23707
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-23707

Query Match 18.4%; Score 451.8; DB 11; Length 483;
Best Local Similarity 99.6%; Pred. No. 5.2e-130;
Matches 453; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1415 CCAGATCCATGAGCTTAACAGGCCAAAGTGAGCACAGTCATGGATACCGTTGGCCGGA 1474
Db 29 CCAGATCCATGAGCTTAACAGGCCAAAGTGAGCACAGTCATGGATACCGTTGGCCGGA 88
QY 1475 GGCCTTCGGGACTCCCTGGGCGGAACCTGTGGACCTGGGCTGACAGGCAACCATGGAGTTG 1534
Db 89 GGCCTTCGGGACTCCCTGGGCGGAACCTGTGGACCTGGGCTGACAGGCAACCATGGAGTTG 148
QY 1535 TGCACCCAGCTTTCGGGACAGTGCCCAACAAGTTTGAGGAGAACACCTACTGCTTCGGCC 1594
Db 149 TGCACCCAGCTTTCGGGACAGTGCCCAACAAGTTTGAGGAGAACACCTACTGCTTCGGCC 208
```

```
QY 1595 GGACCGTGGAGACACTGCTGCTCCGCTTGGCAAGACCATCATGGAGGAGCAGCTGGTAC 1654
Db 209 GGACCGTGGAGACACTGCTGCTCCGCTTGGCAAGACCATCATGGAGGAGCAGCTGGTAC 268
QY 1655 TGAAGCGGTGGCCCAACATCCCTCATCAACCTGTATGGCATGACGGCGCTGCTGTCGCGG 1714
Db 269 TGAAGCGGTGGCCCAACATCCCTCATCAACCTGTATGGCATGACGGCGCTGCTGTCGCGG 328
QY 1715 CCAGCGGCTCCATCCGCAATTCGCGCTCCGCAACACGACGAGCTTCTCTTGGCCAAACA 1774
Db 329 CCAGCGGCTCCATCCGCAATTCGCGCTCCGCAACACGACGAGCTTCTCTTGGCCAAACA 388
QY 1775 CCTTCTCGGTGGAGCTTACTTGCAGAACTCTTTCAGCCTCTCTCAGCTGGACAAGTATG 1834
Db 389 CCTTCTCGGTGGAGCTTACTTGCAGAACTCTCTCAGCCTCTCTCAGCTGGACAAGTATG 448
QY 1835 CTCAGAAAACCTTAGATGACGAGATTAAAGAAAGTGC 1869
Db 449 CTCAGAAAACCTTAGATGACGAGATTAAAGAAAGTGC 483
```

RESULT 8

US-09-918-995-11034
; Sequence 11034, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11034
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-11034

Query Match 15.9%; Score 390.6; DB 11; Length 451;
Best Local Similarity 97.8%; Pred. No. 6.3e-111;
Matches 396; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
QY 651 AGCCACACTAAGTGAAGACATGAAGCACTACATCTCAATGGCTTCCAAAGTCTGGATTAC 710
Db 47 AGCCACACTAAGTGAAGACATGAAGCACTACATCTCAATGGCTTCCAAAGTCTGGATTAC 106
QY 711 TAATGGAGGACTGGCCCAATATTTTACTGTTTGGCAAGACTGAGTCTGTTGATTCTGA 770
Db 107 TAATGGAGGACTGGCCCAATATTTTACTGTTTGGCAAGACTGAGTCTGTTGATTCTGA 166
QY 771 TGGATCAGTGAAGACAAAATACAGCATTCATAGTAGAAGAGACTTTCGTCGAGTCAC 830
Db 167 TGGATCATTGAAGACAAAATACAGCATTCATAGTAGAAGAGACTTTCGTCGAGTCAC 226
QY 831 TAATGGAAACCCGGAAGATAAATAGGCATTTCGGGGCTCCAAACACTTGTGAAGTCCATTT 890
Db 227 TAATGGAAACCCGGAAGATAAATAGGCATTTCGGGGCTCCAAACACTTGTGAAGTCCATTT 286
QY 891 TGAACACCAAGATACCTCTGGAACACATCTTGGAGAGGTCGGAGATGGGTTTAAGGT 950
Db 287 TGAATACCAAGATACCTCTGGAACACATCTTGGAGAGGTCGGAGATGGGTTTAAGGT 346
QY 951 GGCCATGAACATCCTCAACAGCGCGGTTTTCAGCATGGGCGAGCTGCTGGCTGGCTGCT 1010
Db 347 GGTCAATGAACATCCTCAACAGCGCGGTTTTCAGCATGGGCGAGCTGCTGGCTGGCTGCT 406
QY 1011 CAAGAGATTGATTGAATGACTGCTGAGTACGCTGCACAAAGAA 1055
Db 407 CAAGAGATTGATTGAATGACTGCTGAGTACGCTGCACAAAGAA 451
```

RESULT 9

US-09-917-800A-1584
: Sequence 1584, Application US/09917800A
: Patent No. US20020119462A1
: GENERAL INFORMATION:
: APPLICANT: Mendrick, Donna
: APPLICANT: Porter, Mark
: APPLICANT: Johnson, Kory
: APPLICANT: Castle, Arthur
: APPLICANT: Elashoff, Michael
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Molecular Toxicology Modeling
: FILE REFERENCE: 44921-5038-US
: CURRENT APPLICATION NUMBER: US/09/917, 800A
: CURRENT FILING DATE: 2001-07-31
: PRIOR APPLICATION NUMBER: US 60/222,040
: PRIOR FILING DATE: 2000-07-31
: PRIOR APPLICATION NUMBER: US 60/222,880
: PRIOR FILING DATE: 2000-11-02
: PRIOR APPLICATION NUMBER: US 60/290,029
: PRIOR FILING DATE: 2001-05-11
: PRIOR APPLICATION NUMBER: US 60/290,645
: PRIOR FILING DATE: 2001-05-15
: PRIOR APPLICATION NUMBER: US 60/292,336
: PRIOR FILING DATE: 2001-05-22
: PRIOR APPLICATION NUMBER: US 60/295,798
: PRIOR FILING DATE: 2001-06-06
: PRIOR APPLICATION NUMBER: US 60/297,457
: PRIOR FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/298,884
: PRIOR FILING DATE: 2001-06-19
: PRIOR APPLICATION NUMBER: US 60/303,459
: PRIOR FILING DATE: 2001-07-09
: NUMBER OF SEQ ID NOS: 1740
: SOFTWARE: Patent in ver. 2.1
: SEQ ID NO 1584
: LENGTH: 2117
: TYPE: DNA
: ORGANISM: Rattus norvegicus
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012891
: US-09-917-800A-1584

Query Match 15.8%; Score 388.6; DB 10; Length 2117;
Best Local Similarity 54.5%; Pred. No. 7.5e-110;
Matches 891; Conservative 0; Mismatches 729; Indels 15; Gaps 5;

QY	179	CTTTGCGCAAGACGCTTTTCCTAGGCAAAATCAAGAGAAAGAGTTCCTCCATTTCCAG	238
DB	236	CTTTTGCTGTGGGAATGTTCAAAGGCCAGCTTACCACCGCAGGTGTCCCATACCCAT	295
QY	239	AAGTTAGCAAGATGAACCTTAATGAA---ATCAATCAGTTCTTGGGACCCCTGGGAAAT	295
DB	296	CTGTGCTCAATGAAGGACAGACACAAATTTCTCAAGAGCTGTTGGACCACTGGCCGGT	355
QY	296	TCTTCACTGAAGAGTGGACTCCCGGAAATTTGACAGGAGGAAATCCAGATGAAA	355
DB	356	TCTTTGAAGAGTGAATGACCTGCCAAGATGACTCTTGGAGAAGGTGGAGGAGCA	415
QY	356	CTTTGGAGAAATGAAGAGCTAGGCTTTTGGGCTGCAAGTCCCAAGAAATATGTTG	415
DB	416	CTTTGAGGAGCTCAAGAACTTGGGGCATTTGGTCTCAAGTACCAGGAGCTGGGTG	475
QY	416	GCTTGGGTCTTCCCAACCATGTACTCAAGACTAGGGAGATCATCAGCATGATG---	472
DB	476	GTTTGGGCTCTCTAATACCAGTAGCTCGCTTGGCAGAGATTGTGGGCATGATGACC	535
QY	473	GCTCATCACTGTGACCTGGCAGCGCACCGGCTATTGGCTCAAGGGGATCATCTGG	532
DB	536	TTGGTGTAGCGTTACCTGGGAGCCCATCAGAGATCGGTTTCAAAGGCACTCTGTCT	595

QY	533	CTGGCACTGAGAGCAGAAAGCCAAATACTTTGCCCTAAACTGGCTGCCGGGAGCACATTG	592
DB	596	ATGGCACAAAGCCCGAGAGGAAATACCTCCCGAGAGTGGCATCCGGGAGCGGTTGG	655
QY	593	CAGCCTTCTGCTCAGGAGCCAGCAGTGGGAGGATGCGAGCCTCAATCCGGAGCAGAG	652
DB	656	CGGCTTCTGCTGACTGAGCCCTCGAGGGGTGCGATGTGGCCTCTATCCGAAGCTCAG	715
QY	653	CCACACTAAGTGAAGACAGAAGCACTACATCCTCAATGGTCTCAAGGTCTGGATTACTA	712
DB	716	CTGTACCTAGCCCTGTGGAAAGTATTATCTCTCAACGGAAGCAAGATCTGGATCAGTA	775
QY	713	ATGGAGGAGTGGCCAAATATTTTACTGTGTTTTCGAAAGACTGAGTCGTTGATTCTG---	769
DB	776	ATGGGGTCTGGCAGACATTTTCACTGTCTTGGCAAAACGCAATTAAGATGAGGCCA	835
QY	770	ATGGATCACTGAAAGACAAAATCACAGCATTCATAGTAGAAGAGACTTTGTGGAGTCA	829
DB	836	CGGGGCCGTGAAGAGAGATCACAGCTTTCGTAGTGAACGGAGCTTTGGAGGGTTA	895
QY	830	CTAATGGGAAACCCGAAAGATAAATTAGGCATTTGGGGCTTCCAACTCTGTGAAGTCCAT	889
DB	896	CCCATGGCTCCCGAAAAAGAGATGGGCATCAAGGCATCTAACATCATCAGAGGTGACT	955
QY	890	TTCAAAACACACAGATACCTGTGGAAACATCTTGGAGAGTTCGGAGATGGTTTAAGG	949
DB	956	TTGATGGAGTCAAGTGGCCAGCAGAGAATGCTAGGAGAGTGGGAGATGCTTCAAGG	1015
QY	950	TGGCCATGAACATCCTCAACAGCGCGGTTTCAGCATGGGCGAGCGCTGCTGGGCTGC	1009
DB	1016	TTGCTGTCAACATCCTCAACACGGAAGATTTGGGATGGCTGCAACCTTAGCAGCACCA	1075
QY	1010	TCAAGAGATTGATTGAAATGACTGCTGAGTACGCTGCGACAGGAAACAGTTTAAACA	1069
DB	1076	TGAAGCCATCATTTGCCAAGCGGTTGATCTACTACTACCGTACCCAGTTTGGGGACA	1135
QY	1070	GGCTCAGTGAATTTGGATTGATTTCAGGAGAAATTTGCACTGATGGCTCAGAAGGCTT	1129
DB	1136	AAATTCAAACTTTGGGGTATCCAGAAAGCTGCTCGGATGGCTATTCGACGATG	1195
QY	1130	TCATGAGAGTATGACTACCTACAGCAGGAGTGTGGACCACTGCTTCCGCTTCCGACT	1189
DB	1196	TGACTGAGTCCATGGCTTACATGCTGAGTGCCAAACATGGACCCAG---GGATCAAA	1252
QY	1190	GTCCATCGAGCAGCAGCTGTTCAAGGTGTTACGCTCCGAGGCGCTGCGAGTGTGTA	1249
DB	1253	TCAGATAGAGCTGCCATCAGCAAAATCTTGGCTCGGAGGCGGCTGGAAGTGCACAG	1312
QY	1250	GTGAGCGCTGACAGTCTCGGGGCTTGGGCTACACAAAGGAGTATCCGCTAGCAGGCA	1309
DB	1313	ATGATGCTATCCAGATAATGGGGGCTGCTCATGAAGGAACCAAGGCTAGAGCGTG	1372
QY	1310	TACTGGTGCACCCGCTCTCTCTATCTTCGAGGGAACCAATGAGATCTCCGGATGT	1369
DB	1373	TGCTCCGAGATATTCGAATCTTCGGATCTTTGAAGGGAACAAATGACATCTTCGACT	1432
QY	1370	ACATGCGCTTACGCGGCTGACGATGCGCGCGCATCTCTGACTACCAAGATCCATGAGC	1429
DB	1433	TTGTGGCTCTACAAAGGCTGCTGCACAAAGAAAGGAACCTACGGGACTTGGTAATGCC	1492
QY	1430	TTAAACAGGCAAGTGAAGCAGTATGATACCGTTGGCGGAGGCTTCGGGACTCC	1489
DB	1493	TAAAGATCTCTTTGAAATGTTGGCTCTCATAGGAGAAAGCAACACAGCTGAGGC	1552
QY	1490	TGGGCCCACTGTGGACCTGGCTGACAGCAACCATGAGTTGTGACCCCGCTCTTG	1549
DB	1553	GGCGACAGGATTTGGCAGTGGTCTGA---GTCTCTCGGAATTTCCACCCAGAGTTGA	1609
QY	1550	CGGACAGTGGCAACAAGTTTTCAGGAGAACACCTACTGCTTCGCGGAGCGCTGGAGAC	1609
DB	1610	GTCCAGTGTGAACTGGCAGTGCAGCTCTGGAACAATTTGCCACTGTAGTGGAGCGA	1669
QY	1610	TGCTGCTCCGCTTTGGCAAGACCATCATGAGGAGGAGCAGCTGGTACTCAAGCGGCTG	1669

[illegible]

RESULT 10

```

US-09-918-995-16927
; Sequence 16927, Application US/09918995
; Publication NO. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 16927
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-16927

```

[illegible]

RESULT 11

```

US-09-822-849A-480
; Sequence 480, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary

Query Match          7.6%; Score 187.4; DB 11; Length 402;
Best Local Similarity 94.6%; Pred. No. 1.5e-47;
Matches 194; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 11 TCCTCGGCGCTAAGAAGGGGAGACTGTAGGCTGGGGAGCACTGGCGAGCATGA 70
      |||||

```

Db 1 TCCTCCGGCGCTAAGAGGGAGAGTGGAGGCTGTGGTGTGAGACATCTGCCACATGA 60
QY 71 GCGGCTGCGGGCTTCTCTCGCGCACACACCGCTCGCGCTCGTCCCTGCCGGGTCTGGTGG 130
Db 61 GCGGCTGCGGGCTTCTCTCGCGCACACACCGCTCGCGCTCGTCCCTGCCGGGTCTGGTGG 120
QY 131 TCCTACCGGGAACCGGGGCTACTCGCGCACACCGCGCTCTAGCAGCTTTTCGCCAAG 190
Db 121 TATCTACCGCAACCGGGGCTACTCGCGCACCATCCGCTGTAGCAGCTTTTCGCCAAG 180
QY 191 AGCTTTTCCTAGGCAAAATCAAGAA 215
Db 181 AGCTTTTCCTAGGCAAAATCAAGAA 205
RESULT 13
US-09-796-692-8175
; Sequence 8175, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8175
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8175
Query Match 6.7%; Score 165.2; DB 10; Length 568;
Best Local Similarity 58.0%; Pred. No. 1.6e-40;
Matches 312; Conservative 0; Mismatches 223; Indels 3; Gaps 1;
QY 848 ATAATATAGGCAATTCGGGCTCCACACTTGTGAAGTCCATTTTGAACACCAAGATAC 907
Db 3 AGAAGATGGCAATCAAGGCTTCAACACACAGCAGAGGTTCTTGTGATGAGTACGGGTGC 62
QY 908 CTGTGGAACATCTTGGAGAGTCGGAGAGTGGTTTAAAGTGGCCATGAACATCCCTCA 967
Db 63 CATCGGAGAACCTGTGGGTGAGGTTGGAGTGGCTTCAAGTTGCCATGCACATCCCTCA 122
QY 968 ACAGCGCGCGCTTACAGATGGCAGCGCTCGTGGCTGCTCAAGAGATTGATTGAA 1027

Db 123 ACAATGGAAGTTTGGCATGGCTCGGCCCTGGCAGGTACCATGAGAGGCAATTCATTCCTA 182
QY 1028 TGACTGCTGAGTACGCTGCAACAGAAACAGTTTAAACAAGAGCTCAGTGAATTTGGAT 1087
Db 183 AGGCGGTAGATCATGCCACTAATCGTACCCAGTTTGGGAGAAAATTCACAACCTTTGGGC 242
QY 1088 TGATTCAGGAGAAAATTTGCACTGATGGCTCAGAAGGTTAGTTCATCGAGAGTATGACCT 1147
Db 243 TGATCCAGGAGAAGCTGGCAGGATGGTTATGCTGCATATGTAACAGTCCATGCTT 302
QY 1148 ACCTCAGCAGGAGATGCTGACCAACCTGGCTTTCCGAGCTGCTCCATCGAGCAGCA 1207
Db 303 ACATGGTGAGTGTAAACATGGACCA---GGAGGCCAGGACTTCCAGTAGAGCGCGCA 359
QY 1208 TGGTGAAGGTGTTCAAGCTCCGAGCGCGCTGGCAGTGTGTGAGTAGCGCTGCAGATCC 1267
Db 360 TCAGCAAAATCTTTGGCTCGGAGCGAGCTGGAAGGTGACAGATGAATGCATCCAAATCA 419
QY 1268 TCGGGGGCTTGGGCTACACAAGGAGTATCCGTACGAGCGCATACTGGTGCACACCGCA 1327
Db 420 TGGGGGTATGGCTTCATGAAGGAACCTGGAGTAGACGCTGCTCCGAGATCTTCGCA 479
QY 1328 TCCTCCTCATCTTCAGGAGAACCAATGAGATTCTCCGATGTACATCGCCCTGACGGG 1385
Db 480 TCCTCCGATCTTTGAGGGGACAAATGACATTTCTCGCTGTTTGTGGCTCTGCAGGG 537
RESULT 14
US-10-040-862-8175
; Sequence 8175, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-01352085
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8175
; LENGTH: 568

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-8175

Query Match 6.7%; Score 165.2; DB 14; Length 568;
Best Local Similarity 58.0%; Pred. No. 1.6e-40;
Matches 312; Conservative 0; Mismatches 223; Indels 3; Gaps 1;

```
QY 848 ATAAATTAGGATTTCGGGGCTCCAAACACTTGTGAAGTCCATTTTGAACACACCAAGATAC 907
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 AGAAGATGGGCATCAAGGCTTCAACACACAGCAGAGGTGTTCTTTGATGGAGTACGGGTGC 62
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 908 CTGTGGAACACATCCTTGGAGAGTCCGAGATGGGTTTAAAGTGGCCATGAACATCTCA 967
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 CATCGAGAGAGCTGTGGGTGAGTTGGAGTGGCTTCAAGGTTGCCATCCATCCTCA 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 968 ACAGCGCCGGTTCAGCATGGCAGCTGCTGGCTGGCTGGCTGCTCAAGAGATTGATTGAAA 1027
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 ACAATGGAAGTTTGGCATGGCTGCGGCCCTGGCAGGTACCATGAGAGGCATCATTCCTA 182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1028 TGACTGCTGAGTAGCCCTGCACAGGAACAGTTTACAAAGAGGCTCAGTGAATTTGGAT 1087
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 AGCGGTAGATCATGCCACTAATCGTACCCAGTTTGGGGAGAAAATTCACAATTTGGGC 242
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1088 TGATTACAGAGAAATTTGCATGCTCAGAGGCTTACGTCATGGAGAGTATGACCT 1147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 TGATCCAGGAGAGCTGGCAGCGATGTTATGCTGCAGTATGTAACCTGAGTCCATGGCTT 302
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1148 ACCTCAGCAGGAGTGTGGAGCAACCTGGCTTCCCGACTGCTCCATCGAGGAGGCCA 1207
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 ACATGGTGAAGTACATGAGCA ---GGGAGCCAGGACTTCCAGATAGAGGCGGCCA 359
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1208 TGTGAAGGTGTTACGCTCCGAGGCGCCTGGCAGTGTGTGAGTGAGGCGCTGCAGATCC 1267
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 360 TCAGCAAAATCTTTGGCTCGGAGGACGCTGGAAGGTGACAGTGAATCCATCCAAATCA 419
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1268 TCGGGGCTTGGGCTACACAGGACTATCCGTACAGGCGATACGTGCGTGACACCCGCA 1327
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 420 TGGGGGTATGGCTTCATGAAGGAACCTGGAGTGAAGCGTGTGCTCCGAGATCTTCGCA 479
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1328 TCCTCTCATCTTCGAGGAACCAATGAGATTCTCCGGATGTACATCGCCCTGACGGG 1385
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 480 TCTTCGGATCTTTGAGGGACAATGACATTCTTCGGCTGTTTGGCTCTGCAGGG 537
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 15

US-09-867-701-3392

```
; Sequence 3392, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Jones, Paul A.
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3392
; LENGTH: 183
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(183)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-3392
```

Query Match 6.5%; Score 159.4; DB 10; Length 183;
Best Local Similarity 97.7%; Pred. No. 5e-39;
Matches 171; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```
QY 1337 TCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCCCTGACGGGTCTGCAGCATG 1396
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9 TCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCCCTGACGGGTCTGCAGCATG 68
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1397 CCGGCGGCATCTGACTACCGAGATCCATGAGCTTAAACAGGCCAAAGTGAACACAGTCA 1456
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 69 CCGGCGGCATCTGACTACCGAGNTCCATGAGCTTAAACAGGCCAAAGTGAACACAGTCA 128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1457 TGGATACCCGTTGGCGGAGG-CTTCGGGACTCCCTGGGCGCAACTGTGACCTGG 1510
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 129 TGGATACCCGTTGGCGGAGGTTTTCGGGACTCCCTGGGCGCAACTGTGACCTGG 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: September 7, 2003, 05:58:53
Job time : 563.612 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: September 6, 2003, 23:10:50 ; Search time 5159.14 Seconds
(without alignments)
11551.252 Million cell updates/sec

Title: US-09-945-326-1
Perfect score: 2452
Sequence: 1 cgtgtgtgttccttcgccc.....actgttaaaaaaaaaaaaaa 2452

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estlin:*
 - 4: em_estmu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_hic:*
 - 9: gb_est1:*
 - 10: gb_est2:*
 - 11: gb_hic:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: em_gss_hum:*
 - 18: em_gss_inv:*
 - 19: em_gss_pln:*
 - 20: em_gss_vrt:*
 - 21: em_gss_fun:*
 - 22: em_gss_nam:*
 - 23: em_gss_mus:*
 - 24: em_gss_pro:*
 - 25: em_gss_rod:*
 - 26: em_gss_phg:*
 - 27: em_gss_vrl:*
 - 28: gb_gss1:*
 - 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1683.6	68.7	1914	11 BC041572	BC041572 Homo sapi
2	1382	56.4	2432	11 AK075984	AK075984 Mus muscu
3	1379.4	56.3	2446	11 AK031820	AK031820 Mus muscu
4	1082.2	44.1	2906	11 AK028571	AK028571 Mus muscu

5	1026	41.8	1123	11 AF078854	AF078854 Homo sapi
6	952.8	38.9	1071	12 BM080820	BM080820 AGENCOURT
7	940.4	38.4	1201	9 AL539220	AL539220 AL539220
8	937.8	38.2	1201	13 BX458691	BX458691 BX458691
9	879.2	35.9	1201	9 AL567735	AL567735 AL567735
10	869.6	35.5	969	9 AL549834	AL549834 AL549834
11	867.2	35.4	1010	12 BM552756	BM552756 AGENCOURT
12	859.4	35.0	1079	12 BM561207	BM561207 AGENCOURT
13	854	34.8	1206	9 AL582366	AL582366 AL582366
14	852.4	34.8	890	13 BQ938039	BQ938039 AGENCOURT
15	852.4	34.8	1201	9 AL560956	AL560956 AL560956
16	847.2	34.6	905	13 BQ541780	BQ541780 AGENCOURT
17	844.4	34.4	1020	13 BQ196321	BQ196321 AGENCOURT
18	838.6	34.2	1186	9 AL578429	AL578429 AGENCOURT
19	833.2	34.0	914	9 AL524997	AL524997 AL524997
20	822.2	33.5	1023	13 BQ072180	BQ072180 AGENCOURT
21	813.4	33.2	851	13 BQ176038	BQ176038 AGENCOURT
22	810.6	33.1	884	14 CA488487	CA488487 AGENCOURT
23	810	33.0	922	13 BQ528453	BQ528453 AGENCOURT
24	809.6	33.0	1199	9 AL556077	AL556077 AL556077
25	808	33.0	831	12 B1524078	B1524078 603052267
26	807.8	32.9	1023	12 BM0809053	BM0809053 AGENCOURT
27	805.2	32.8	875	13 BQ882994	BQ882994 AGENCOURT
28	799.6	32.6	885	13 BQ191988	BQ191988 AGENCOURT
29	798.8	32.6	888	13 BQ960770	BQ960770 AGENCOURT
30	798.2	32.6	848	9 AL557534	AL557534 AL557534
31	797.6	32.5	937	13 BQ279135	BQ279135 AGENCOURT
32	793.2	32.3	798	12 B1762564	B1762564 603048696
33	792	32.3	980	12 BQ054246	BQ054246 AGENCOURT
34	790.4	32.2	857	9 AL557535	AL557535 AL557535
35	788.4	32.0	867	10 BG674212	BG674212 602619872
36	784.4	32.0	799	12 B1914477	B1914477 603182269
37	780.2	31.8	891	13 BX328269	BX328269 BX328269
38	772.4	31.5	851	10 BG757623	BG757623 602711206
39	767.4	31.3	885	13 BQ156414	BQ156414 AGENCOURT
40	756	30.8	905	9 AL524996	AL524996 AL524996
41	755.6	30.8	893	12 B1772602	B1772602 603033154
42	754.2	30.8	973	10 BE792118	BE792118 601581681
43	750.6	30.6	771	9 A0142885	A0142885 A0142885
44	748.8	30.5	761	12 B1821212	B1821212 603034922
45	747.4	30.5	784	12 B1829609	B1829609 603079347

ALIGNMENTS

RESULT 1
BC041572

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

BC041572 Homo sapiens, Similar to NP0002 protein, clone IMAGE:3952086, mRNA.
BC041572
BC041572.1 GI:27469662

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1914)

Strausberg, R.

Direct Submission

Submitted (20-DEC-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

Contact: MGC help desk

Email: cgapsb-remail.nih.gov

Tissue Procurement: DCTD/DP

cDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

REMARK

COMMENT

Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgen@nri.nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, G.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 44 Row: c Column: 14
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis

This clone has the following problem: frame shifted.

FEATURES

source	Location/Qualifiers
	1..1914
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="IMAGE:3952086"
	/tissue_type="Ovary, adenocarcinoma"
	/clone_lib="NIH_MGC_9"
	/lab_host="DH10B-R"
	/note="vector: pOTB7"
BASE COUNT	477 a 482 c 522 g 433 t
ORIGIN	
Query Match	68.7%; Score 1683.6; DB 11; Length 1914;
Best Local Similarity	97.7%; Pred. No. 0;
Matches 1741, Conservative	0; Mismatches 9; Indels 32; Gaps 2;
QY	698 AGGTCGGATTACTAATGGAGGACTGCCAATATTTTACTGCTTTGCAAGACGTAGG 757
DB	133 AGGTCGGATTACTAATGGAGGACTGCCAATATTTTACTGCTTTGCAAGACGTAGG 192
QY	758 TCGTGTGATTCGATGGATCAGTGAAGACAAATATCAGACATTCATAGTAGAAGAGACT 817
DB	193 TCGTGTGATTCGATGGATCAGTGAAGACAAATATCAGACATTCATAGTAGAAGAGACT 252
QY	818 TTGTTGGAGTCACTAATGGGAAACCCGAAGATAAATAGGATTCGGGGCTCCAACTT 877
DB	253 TTGTTGGAGTCACTAATGGGAAACCCGAAGATAAATAGGATTCGGGGCTCCAACTT 312
QY	878 GTGAAGTCCATTGAAACACCAAGATACCTGTGGAACATCCTTTGGAGAGTCGGAG 937
DB	313 GTGAAGTCCATTGAAACACCAAGATACCTGTGGAACATCCTTTGGAGAGTCGGAG 372
QY	938 ATGGGTTTAAAGTGGCCATGAACATCCCTCAACAGCGCGGTTTCAGCATGGGAGTCG 997
DB	373 ATGGGTTTAAAGTGGCCATGAACATCCCTCAACAGCGCGGTTTCAGCATGGGAGTCG 432
QY	998 TGGCTGGGCTGCTCAAGAGATGATT-----GAAATGACTGCTGAGTACGCTGCACAAAG 1053
DB	433 TGGCTGGGCTGCTCAAGAGATGATTGTAAGATGACTGCTGAGTACGCTGCACAAAG 492
QY	1054 AAACAGTTTAAACAGAGGCTCAGTGAATTTGGATGATT----- 1092
DB	493 AAACAGTTTAAACAGAGGCTCAGTGAATTTGGATGATTGCTTTGTTTCTG 552
QY	1093 -----CAGGAGAAATTTGCACTGATGGCTCAGAGGCTTACGTCATGGAGATGAC 1145
DB	553 AAACACTCCAGAGAAATTTGCACTGATGGCTCAGAGGCTTACGTCATGGAGATGAC 612
QY	1146 CTACCTCACAGGAGGATGCTGGACCAACCTGGCTTTCCGACTGCTCCATCGAGGAGC 1205
DB	613 CTACCTCACAGGAGGATGCTGGACCAACCTGGCTTTCCGACTGCTCCATCGAGGAGC 672
QY	1206 CATGTTGAAGGTTTACGCTCCGAGCGCCCTGGCAGTGTGTGAGTGGAGGCTGCAGAT 1265

Db	673	CATGGTGAAGGTGTTACAGTCCGAGGCGCCCTGGCAGTGTGTGAGTGGCGCTGCAGAT	732
QY	1266	CCTCGGGGCTTGGGCTACACAAGGGACTATCCGTACGAGCGCATCTGGTGACACCCG	1325
Db	733	CCTCGGGGCTTGGGCTACACAAGGGACTATCCGTACGAGCGCATCTGGTGACACCCG	792
QY	1326	CATCCTCCTCATCTTCGAGGAAACCAATGAGATTCCTCCGATGTACATCCCTGACGGG	1385
Db	793	CATCCTCCTCATCTTCGAGGAAACCAATGAGATTCCTCCGATGTACATCCCTGACGGG	852
QY	1386	TCTCAGCATGCGCGCGCATCTCTGACTACCAAGGATCCATGAGCTTAAACAGGCCAAAGT	1445
Db	853	TCTCAGCATGCGCGCGCATCTCTGACTACCAAGGATCCATGAGCTTAAACAGGCCAAAGT	912
QY	1446	GAGCAGATCATGATACCGTTGGCGGAGGCTTGGGAGTCCCTGGCGGAGACTGTGGA	1505
Db	913	GAGCAGATCATGATACCGTTGGCGGAGGCTTGGGAGTCCCTGGCGGAGACTGTGGA	972
QY	1506	CTGGGGCTGACAGCAACCATGAGTGTGACCCCGAGTCTTGGGAGAGTGCACAACAA	1565
Db	973	CTGGGGCTGACAGCAACCATGAGTGTGACCCCGAGTCTTGGGAGAGTGCACAACAA	1032
QY	1566	GTTCGAGGAGAACCTTACTCTTGGCGGAGGCTGGGAGACACTGCTGCTCCGCTTTGG	1625
Db	1033	GTTCGAGGAGAACCTTACTCTTGGCGGAGGCTGGGAGACACTGCTGCTCCGCTTTGG	1092
QY	1626	CAAGACCATCATGAGGAGGAGTGTGACTGAAGGGGTGGCGGAGTCCCTCATCAACCT	1685
Db	1093	CAAGACCATCATGAGGAGGAGTGTGACTGAAGGGGTGGCGGAGTCCCTCATCAACCT	1152
QY	1686	GTATGGCATGACGCGCGTGTCTCGCGGCGGAGGCTCCATCCGCAATTTGGGCTCCGAA	1745
Db	1153	GTATGGCATGACGCGCGTGTCTCGCGGCGGAGGCTCCATCCGCAATTTGGGCTCCGAA	1212
QY	1746	CCAGCAGCACAGGTTCTTGGCCACACCTTCTGCGTGGAGGCTTACTTGCAGAACT	1805
Db	1213	CCAGCAGCACAGGTTCTTGGCCACACCTTCTGCGTGGAGGCTTACTTGCAGAACT	1272
QY	1806	CTTCAGCTCTCTCAGCTGGACAACTATGCTCCAGAAAACCTTAGATGACAGATTAAGAA	1865
Db	1273	CTTCAGCTCTCTCAGCTGGACAACTATGCTCCAGAAAACCTTAGATGACAGATTAAGAA	1332
QY	1866	AGTGTCCAGCAGATCTTGTGAGAGGAGGCTTATCTGTGCGCCACCTCTGCAGAGAC	1925
Db	1333	AGTGTCCAGCAGATCTTGTGAGAGGAGGCTTATCTGTGCGCCACCTCTGCAGAGAC	1392
QY	1926	ATGCTGAGCAGGAGGAGTGTCCCTGCTACCGCGCGGCTTACCCATGGCCCGTGTCT	1985
Db	1393	ATGCTGAGCAGGAGGAGTGTCCCTGCTACCGCGCGGCTTACCCATGGCCCGTGTCT	1452
QY	1986	GGATGACTGTTACTCTTTTTCAGAAAGTGTGGGATATATCAGAGTTAAGCTTTTGT	2045
Db	1453	GGATGACTGTTACTCTTTTTCAGAAAGTGTGGGATATATCAGAGTTAAGCTTTTGT	1512
QY	2046	CCCGCTGTCACCTGAGGTTGTCCCTGCGCTGGGAGGCTCTTCCAGAGTTTGTACC	2105
Db	1513	CCCGCTGTCACCTGAGGTTGTCCCTGCGCTGGGAGGCTCTTCCAGAGTTTGTACC	1572
QY	2106	TGAGGAGGAGTGTCTCTTAACAGGACCATCACAGCTTCTGAAGTGGCGGAGAGAGAA	2165
Db	1573	TGAGGAGGAGTGTCTCTTAACAGGACCATCACAGCTTCTGAAGTGGCGGAGAGAGAA	1632
QY	2166	TGGAATGCTGACCCCTGGAACTGGCGGGTATTTCTGTCATTGAGGAGACCATAGTG	2225
Db	1633	TGGAATGCTGACCCCTGGAACTGGCGGGTATTTCTGTCATTGAGGAGACCATAGTG	1692
QY	2226	AACTGGGCTTATGCTGCTGCCCTCCAGGCTGTGAGGTGGGTGGGACCTTGTCTCAGGT	2285
Db	1693	AACTGGGCTTATGCTGCTGCCCTCCAGGCTGTGAGGTGGGTGGGACCTTGTCTCAGGT	1752
QY	2286	TGGATAGCCATTTCTGCTCAACACACATTTCTTAAGAAACAGCTTGAAGCTCTGCTG	2345
Db	1753	TGGATAGCCATTTCTGCTCAACACACATTTCTTAAGAAACAGCTTGAAGCTCTGCTG	1812

QY 2346 GGTCAATTCATTAACTAGACAGCAGCCACTTAAACATGTCACGAGCAACCATTTAAACA 2405
 |||||
 Db 1813 GGTCAATTCATTAACTAGACAGCAGCCACTTAAACATGTCACGAGCAACCATTTAAACA 1872
 |||||
 QY 2406 AAGATATATAAATGTCACAACTCTGTGACTGTTAAAAA 2447
 |||||
 Db 1873 AAGATATATAAATGTCACAACTCTGAAAAA 1914
 |||||

RESULT 2
 Locus AK075984
 DEFINITION Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
 enriched library, clone:2600017P15 product:VERY-LONG-CHAIN ACYL-COA
 DEHYDROGENASE VLCAD homolog [Homo sapiens], full insert sequence.
 ACCESSION AK075984
 VERSION 1 GI:26344893
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 JOURNAL Sequencing pipeline with 384 multipeptide sequencer
 MEDLINE 20530913
 PUBMED 120530913

REFERENCE
 AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamahata, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
 Quackenbush, J., Schriml, L.M., Stauber, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
 Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S.,
 and Hayashizaki, Y.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409 (6821), 685-690 (2001)
 MEDLINE 21085660

PUBMED
REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE
AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

BASE COUNT
 ORIGIN

Query Match

11217851

5

The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

6 (bases 1 to 2432)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Araki, A.,
 Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
 Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
 Itoh, M., Itoh, Y., Ito, Y., Ito, Y., Ito, Y., Ito, Y., Ito, Y.,
 Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Konda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyezaki, A., Nishi, K.,
 Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
 Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
 Shibata, K., Shibata, Y., Shihagawa, A., Shiraki, T., Sogabe, Y.,
 Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
 Tejima, Y., Toya, T., Yamamura, T., Yamahata, I., Yasunishi, A.,
 Yoshida, K., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

Submitted (16-APR-2002)

Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

Location/Qualifiers

1..2432

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM,DB:2600017P15"

/db_xref="MGI:1895339"

/db_xref="taxon:10090"

/clone="2600017P15"

/tissue_type="whole body"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="10 days embryo"

29..1906

/note="unnamed protein product; VERY-LONG-CHAIN ACYL-COA
 DEHYDROGENASE VLCAD homolog [Homo sapiens] (SPTR1AAL56011,
 evidence: FASTV, 85.6%id, 99.6%length, match=1869)
 putative"

/codon_start=1

/protein_id="BAC36096.1"

/db_xref="GI:26344894"

/db_xref="MGI:1914272"

/translation="MSGCVLLSRGATAAARASRVLRREFARRRPLHTLSQCSFA
 KELFGLNIEQKGVFPPEVYQHELSINQFVPLEKFEEDVSRKIDQEGKIPVDTL
 EKLSGLGFLQVPEEYGGGLSNTMYARLGEIISLDSITVTLAAHQIGUKGILV
 GNEQKALPLKLSGEHIAAFCLTPEPASGDAASIQTRATLSDEKQVFIILGSKVIT
 TNGGLANFTVFAKTEVVDSDGKTKMTAFIVERDFGTTNGKPEKLGIRSGNTE
 VHFENRTVENVLGEVGGFKVANNILNSGRFSNGSAVAGNKKLIELTAYACTRK
 QFNRLSEFGLIOEKFAALMAQKAYMESMYLTSGMLDQPGFPCDSEAAVMKVFSE
 AAQVCSALQILGSGFMKDIPTERMURDARILLIFEGTNEILRLFTALTGLQHGR
 ILTSRIKSLKSGNVTVMETGRKLRSLDGLTGLDGVVHPSRSLKLENN
 VHYFGRTVETLLRFGKNIVEEQLVRANILNLYGMTAVLSRASRSTRIGLRNHD
 HEVLLANNFCVEAYFQNLFSQLDKNAPENLDFQIKKVSQIILEKRAYICANPLDRA
 S"

640 a 564 c 652 g 576 t

56.4%; Score 1382; DB 11; Length 2432;

Best Local Similarity 84.3%; Pred. No. 1.6e-292;			Matches 1556; Conservative 0; Mismatches 290; Indels 0; Gaps 0;		
QY	98	CGGCTGCGGCTCGTGCCTGCGGGTGTGGTGTCTACCGCGAACCGGGCTACTGCG	157		
Db	72	CGGCTCGGGGAGGTTCCCGGGTGTGAGGAATTCACGCGAGGCGCGCGGTGCG	131		
QY	158	GCACACCGCGGCTGTACGAGCTTTGCGCAAGAGCTTTTCTAGGCAATCAAGAAGA	217		
Db	132	ACACACGCTGCGAGAGCTGCTGGTTCGCAAGAGCTTCTCTGGGCAACATCGAGCAGA	191		
QY	218	AGAAATTTCCCATTTCCAGAGTTAGCCAGATGAACCTTAATGAATCAATCACTTCT	277		
Db	192	AAGGAGTTTCCCATTTCCAGAGTTAGCCAGATGAACCTTAATGAATCAATCACTTCT	251		
QY	278	TGGGACCGTGGAAAAATTTCTCACTGAAGAGGTGGACTCCGCAAAATTTGACAGGAG	337		
Db	252	TGGGACCGTGGAAAAATTTCTCACTGAAGAGGTGGACTCCGCAAAATTTGACAGGAG	311		
QY	338	GGAAATTTCCCATTTCCAGAGTTAGCCAGATGAACCTTAATGAATCAATCACTTCT	397		
Db	312	GAATAATTTCCCATTTCCAGAGTTAGCCAGATGAACCTTAATGAATCAATCACTTCT	371		
QY	398	TCCAGAGAAATATGGTGGCTGGGCTTCTCAACACCATGATGACTCAAGACTAGGGAGA	457		
Db	372	TCCAGAGAAATATGGTGGCTGGGCTTCTCAACACCATGATGATGCTGGCTCGGGAGA	431		
QY	458	TCATCAGCATGGATGGGTGCATCACTGTGACCTTGGCAGCGCACAGGCTATTGGCTCA	517		
Db	432	TCATCAGCATGGATGGGTGCATCACTGTGACCTTGGCAGCGCACAGGCTATTGGCTCA	491		
QY	518	AGGGGATCATCTGGCTGGCACTGAGGAGCAGAAAGCAATATCTTGCCTAACTGGCGT	577		
Db	492	AGGGGATCATCTGGCTGGCACTGAGGAGCAGAAAGCAATATCTTGCCTAACTGGCGT	551		
QY	578	CCGGGAGCAGCATTTGAGCCTTTGCTCAGGAGCAGCCAGTGGGAGGCTGACGCT	637		
Db	552	CTGGGAGCAGCATTTGAGCCTTTGCTCAGGAGCAGCCAGTGGGAGGCTGACGCT	611		
QY	638	CAATCCGAGCAGAGCCACACTAAGTGAAGACAGAAAGCAATCACTCCTCAATGGCTCA	697		
Db	612	CAATCCGAGCAGAGCCACACTAAGTGAAGACAGAAAGCAATCACTCCTCAATGGCTCA	671		
QY	698	AGGCTGGATTAATTTGAGGAGTGGGCAATATTTTACTGTGTTTGAAAGACTGAGG	757		
Db	672	AGGTTGGATCACCAGTGGAGGCTGGGCAATATTTTACTGTGTTTGAAAGACTGAGG	731		
QY	758	TCGTTGATCTTGATGATCAGTGAAGACAAATTCAGAGATTCATAGTAGAAAGAGACT	817		
Db	732	TGGTGCATCTGATGGTTTGAAGACAGCAAAATGACCGCATTCATAGTAGAAAGAGACT	791		
QY	818	TTGTTGGAGTCACTAATGGGAAACCCGGAAGATAAATAGGATTCGGGGCTCCACACTT	877		
Db	792	TGGGCGGAATCACTAATGGGAAACCCGGAAGATAAATAGGATTCGGAGGCTCCACACAT	851		
QY	878	GTGAGTCCATTTGAAACACACAGATACCTGTGAAACACATCTTGGAGAGTCCGAG	937		
Db	852	GTGAGTCCATTTGAAACACACAGATACCTGTGAAACACATCTTGGAGAGTTCGAG	911		
QY	938	ATGGGTTTAAAGTGGCCATGAACATCTTCAACAGCGCGGTTGAGCATGGGAGCTGCG	997		
Db	912	GGGGCTTTAAGTACCATGAACATCTTCAACAGCGCGGTTGAGCATGGGAGCTGCGT	971		
QY	998	TGGCTGGGCTCAAGAGATGATTAATGAGTACTGCTGAGTACGCTGCAACAGGAAAC	1057		
Db	972	TGGCGGGATGCTGAAGAGCTGATCGAACTGACTGCTGAGTATGCCCTGTACGAGAAAC	1031		
QY	1058	AGTTTAAACAGAGCTCAGTGAATTTGGATTCAGGAGAAATTTGCATGTATGGCTC	1117		
Db	1032	AGTTTAAACAGAGCTCAGTGAATTTGGATTCAGGAGAAATTTGCATGTATGGCTC	1091		
QY	1118	AGAAGGCTTACGTCATGGAGAGTATGACCTACCTCAGCAGGAGGATGCTGACCAACCTG	1177		
Db	1092	AGAAGGCTTACGTCATGGAGAGTATGACCTACCTCAGCAGGAGGATGCTGACCAACAG	1151		
QY	1178	GCATTCCCGAGCTGCTCCATCAGCAGCCATGTTGAAGTGTTCAGTCCCGAGGCGGCT	1237		
Db	1152	GATTTCGCGAGCTGCTCTATCGAGCGCATGTTGAAGTGTTCAGTCCCGAGGCTGCT	1211		
QY	1238	GGCAGTGTGTGAGTGGGCGCTGCAGATCTCTGGGGCTTGGGCTACACAAAGGACTATC	1297		
Db	1212	GGCAGTGTGTGAGTGGGCGCTGCAGATCTCTGGGGCTTGGGCTACACAAAGGACTATC	1271		
QY	1298	CGTACGAGCGCATCTGCGTGCACACCGCATCTCTCTCATCTTCGAGGGAACCAATGAGA	1357		
Db	1272	CCTATGAGCGCATCTGCGTGCACACCGCATCTCTCTCATCTTCGAGGGAACCAATGAGA	1331		
QY	1358	TTCTCCGATGATGATGCGGCTGACGAGCTGTCAGCATGCGGCGGCTGCTGACTACCA	1417		
Db	1332	TTCTCCGATGATGATGCGGCTGACGAGCTGTCAGCATGCGGCGGCTGCTGACTACCA	1391		
QY	1418	GGATCCATGAGCTTAAACAGCGCAAGTGGACGACGACGATGATGATGCTTGGCGGAGGC	1477		
Db	1392	GGATCCATGAGCTTAAACAGCGCAAGTGGACGACGACGATGATGATGCTTGGCGGAGGC	1451		
QY	1478	TTGGGAGCTCCCTGGGCGGAACTGTGGACCTTGGGCGGCTAACAGGCGATCTTGGTGTGTC	1537		
Db	1452	TTGGGAGCTCCCTGGGCGGAACTGTGGACCTTGGGCGGCTAACAGGCGATCTTGGTGTGTC	1511		
QY	1538	ACCCGAGCTTGGGAGAGTGGCAACAGTTTGGAGGAACACCTACTGCTTCGCGCGCA	1597		
Db	1512	ACCCGAGCTTGGGAGAGTGGCAACAGTTTGGAGGAACACCTACTGCTTCGCGCGCA	1571		
QY	1598	CGGTGAGACACTGCTGCTCCGCTTGGCAAGACCATCATGAGGAGAGCTGTTGACTGA	1657		
Db	1572	CGGTGAGACACTGCTGCTCCGCTTGGCAAGACCATCATGAGGAGAGCTGTTGACTGA	1631		
QY	1658	ACCGGCTGCGCAACATCTCATCACTGATGATGATGATGATGATGATGATGATGATGATG	1717		
Db	1632	ACCGGCTGCGCAACATCTCATCACTGATGATGATGATGATGATGATGATGATGATGATG	1691		
QY	1718	CGCGTCCATCCGCTTGGGCTCCGCAACACGACGACGAGGCTTCTTGGCCCAACACT	1777		
Db	1692	CGCGTCCATCCGCTTGGGCTCCGCAACACGACGACGAGGCTTCTTGGCCCAACACT	1751		
QY	1778	TCGCTGGAGGCTTACTTGGAGATCTCTTCAGGCTCTCTCAGCTGGACAGATGCTC	1837		
Db	1752	TCGCTGGAGGCTTACTTGGAGATCTCTTCAGGCTCTCTCAGCTGGACAGATGCTC	1811		
QY	1838	CAGAAACCTGATGATGAGGATTAAGAAAGTGTCCAGCAGATCTTGGAGAGCGAGCT	1897		
Db	1812	CAGAAACCTGATGATGAGGATTAAGAAAGTGTCCAGCAGATCTTGGAGAGCGAGCT	1871		
QY	1898	ATATCTGTGCCCCCTCTGGAGAGATGCTGAGGAGGAGCA	1943		
Db	1872	ATATCTGTGCCCCCTCTGGAGAGATGCTGAGGCTGGAACCA	1917		
RESULT 3					
AK031820					
LOCUS					
DEFINITION					
Mus musculus adult male medulla oblongata cDNA, RIKEN full-length					
enriched library, clone:630407H09 product:VERY-LONG-CHAIN ACTL-COA					
DEHYDROGENASE VLCD homolog [Homo sapiens], full insert sequence.					
AK031820					
GI:26327642					
VERSION					
AK031820.1					
KEYWORDS					
HTC; CAP trapper.					
SOURCE					
Mus musculus					
ORGANISM					
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
REFERENCE					
1					
Carninci,P. and Hayashizaki,Y.					
High-efficiency full-length cDNA cloning					
JOURNAL					
Meth. Enzymol. 303, 19-44 (1999)					
MEDLINE					
99279253					

2446 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult male medulla oblongata cDNA, RIKEN full-length
enriched library, clone:630407H09 product:VERY-LONG-CHAIN ACTL-COA
DEHYDROGENASE VLCD homolog [Homo sapiens], full insert sequence.

PUBMED

REFERENCE
AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)

20499374

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunuma, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staib, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, K., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombert, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

21085660

11217851

5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 2446)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takehashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

TITLE

JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

FEATURES

source

1. 2446

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:6330407H09"

/db_xref="taxon:10090"

/clone="6330407H09"

/sex="male"

/tissue_type="medulla oblongata"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

42. 1919

/note="unnamed protein product: VERY-LONG-CHAIN ACYL-CoA DEHYDROGENASE VL-CAD homolog [Homo sapiens] (SPTRIAL56011, evidence: FASTY, 85.6%ID, 99.6%length, match=1869)

putative"

/codon_start=1

/protein_id="BAC27565.1"

/db_xref="GI:26327643"

/translation="MSGCVLLSRGATAAGAAASRVLRFTARRPLHTSIQSCSFA KFLGNTBQKGVFPPEVSHSELSINQFVGPLEKFEEDVDSKDQEKIPVDTL EKLGLGFGIOVPEYGGIGLSNPMYARLGEIISLDASITVTLANHQAIGLGLV GNEOKAKYLPKLSGEGHIAFCLPASPAGSAAISQTRATLSDPKFKYLINGSKWI TNGGLANFTVFAKTEVDSDCKMTAFIVDFGGITNGKPLDIEATRYACTRK QFNRLSEFLQERFALMAQKAYVMSNAYLITSCMLDQPGFPCDSIAAAVVFSS EAAQCWSEALQIGSGYMKDYPERMLRDARILLIFEGTNEILRLFTALGLOHAG RILTSRIKELGNSNTVMTIGRLRDSIGRTVDLGTGLGVVHPVSGDSANKLEEN VHYFGRTVETLLRGKNIQVEQLVKRVANVLNLYGNTAVLSRASRSIRIGLNHHD HVLNANPCVEAYFQNLFSLSQDKNAPENLDPOIKKVSQIILEKRAYICAHPLDRA S"

BASE COUNT 641 a 567 c 660 g 578 t

ORIGIN

Query Match

Best 4-local Similarity 56.3%; Score 1379.4; DB 11; Length 2446;

Matches 1554; Conservative 0; Mismatches 291; Indels 0; Gaps 0;

QY

99 GCCTCGGGCTGCTGCTCGGGGGTCTGGTGGTCTCTACCGCGAACCGCGGCTACTCGC 158

Db 86 GCCTCGGGCGGAGCTTCCCGGGTCTGAGGGAATTCACGGCGAGCGCGCCGCTGCA 145

QY 159 CACCAGCCGCTGCTACGAGCTTTCGCCAAGAGAGCTTTCTCTAGCGAAATCAAGAGAA 218

Db 146 CACCAGCTCGAGAGCTGCTCGTTCGCCAAGAGAGCTTCTCTGCGCAACATCGAGCAGAA 205

QY 219 AGAAGTTTTCATTTCCAGAAAGTAGCAGAGATGAATTAATCAATCAATCAATCAATCAAT 278

Db 206 AGGAGTTTTCATTTCCAGAGGTCAGTCAACATGAATCTAGTCAATCAATCAATCAATCAAT 265

QY 279 GGGACCCGTGAAAAATTTCTCACTGAAGAGGTGGACTCCCGAAAAAATTCACCAAGAGG 338

Db 266 GGGACCCGTGAAAAATTTCTCACTGAAGAGGTGGACTCTCGAAAAAATTCACCAAGAGG 325

QY 339 GAAATCCAGATCAAACTTTTGAGAAATTTGAAGAGCTAGGCGCTTTTTCGGCTGCAAGT 398

Db 326 AAAATCCAGTTTTCACACCTTAGAGAGAGTTGAAGAGCTCGGACTTTTTCGGCATACAGGT 385

QY 399 CCGAGAGATATGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 458

Db 386 CCGAGAGATATGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 445

QY 459 CATCAGCATGGTGGTCCATCATCTGACCTTGGAGCCCTGGAGCCGACAGCTATTCGCTCAA 518

Db	446	CATCAGCCTGGATGCCTCTATACGGTGACCGCTGGCAGCACACCGGCTATCGGCTCAA	505
Qy	519	GGGAGCATCTTGGCTGGCACTGAGGAGCAGAAAGCCAAATACTTGCCTAAACTGGCGTC	578
Db	506	GGGATATCTTGGTCGGCAATGAGGAACAGAAAGGCGCAAGTATCTGCCCAAACTGTCGTC	565
Qy	579	CGGGAGCAGCATTCAGCGCTTCTGCCTCAGGAGCAGCCAGCAGGAGGAGGATGACGCCNC	638
Db	566	TGGGAACACATCGCAGCCTTCTGCCTAACAGAGCCAGCGAGTGGGAGTGATGCTCGGTC	625
Qy	639	AATCCGAGCAGACACACTAAGTGAAGACAAGAGCAGCTACATCCTCAATGGCTCAA	698
Db	626	CATCCAGCAGCAGCTACGTTAAGTGAAGATAAGAGTACTTCATCTCAATGGCTCAA	685
Qy	699	GGTCTGGATCTAATGGAGACTGGCCAAATATTTTACTGTGTTTGCRAAGACTGAGGT	758
Db	686	GGTTTGGATCACCNAATGGAGACTGGCCAAATATTTTACTGTGTTTGCRAAAACCGAGGT	745
Qy	759	CGTTGATCTGATGCAGTGAAGAGACAAATCACAGCATTCATAGTAGAAGAGACTT	818
Db	746	GGTCGATCTGATGTTGGAACACAGACAAATGACCGCATTCATAGTAGAAGAGACTT	805
Qy	819	TGGTGGAGTCTACTAATGGGAACCCGAGATAAATTAGGCATTCGGGGCTCCACACTTG	878
Db	806	CGGGGAATCACTAATGGGAACCTGAAGATAAATTAGGCATTCGAGGCTCCACACATG	865
Qy	879	TGAAGTCCATTTTGAACACCAAGATACCTGTGGAACACATCCTTGGAGAGTTCGGAGA	938
Db	866	TGAGTCCATTTTGAANAATACAGAGTCCCTGTCGGAANAATGCTTGGAGAAGTGGAGG	925
Qy	939	TGGGTTTAAGTGGCCATGAACATCCTCAACAGCGCGGTTTCAGCATGGGCGCTGCT	998
Db	926	GGGCTTTAAGTGAAGCATGAACATCCTGAACAGTGGCGATTTCAGCATGGGCGCTGCT	985
Qy	999	GGCTGGGCTCTCAAGAGATTTGATTAATGACATGCTGAGTACCGCTGCACAAGAAACA	1058
Db	986	GGCCGGGATGCTGAAGAGAGCTGACGAAGTCTGAGTACCGAGGATTCAGCATGGGCGCTGCT	1045
Qy	1059	GTTTAAACAGAGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGACATGATGGCTCA	1118
Db	1046	GTTCAACAGGATCTCAGTGAATTTGGTTGATCCAGGAAAGTTTGGCTCATGGCTCA	1105
Qy	1119	GAAGGCTTACGTCATGGAGATGATGACCTACTCACAGCAGGATGCTGGACCAACCTGG	1178
Db	1106	GAAGGCATACGTAATGGAGATGATGGCTTACCTCACTTCAAGGATGCTAGACCAACAGG	1165
Qy	1179	CTTTCGCGACTCTCCATCGAGGAGCCATGGTGAAGTGTTCAGCTCCGAGGCGGCTG	1238
Db	1166	ATTTCGCGACTCTCTATCGAGGAGCCATGCTGAAGTGTTCAGCTCTGAGGCTGCTG	1225
Qy	1239	CGAGTGTGAGTGGGCGCTGCAGATCCTCGGGGCTTGGGCTACACAAGGAGCTATCC	1298
Db	1226	CGAGTGTGAGGAGGCTCTGCAGATCCTTGGGGCTCAGGCTACATGAAGGACTACCC	1285
Qy	1299	GTACGAGCCATCTGCTGACCGGAGCCGATCCTCTCATCTTCGAGGGAACCAATGAGAT	1358
Db	1286	CTATGAGCGCATGCTGCGTGAATGCGCGCATTCCTTAATCTTTGAGGGAACCAATGAGAT	1345
Qy	1359	TCTCCGGATGATACATCGCCCTGACGGGTGTCAGCATGCGCGCCCATCTCTGACTACGAG	1418
Db	1346	TCTTCGGTGTTCATTTGCCCTGACAGGCGCTGCAGCATGCTGACGCACTCTGACCTCGAG	1405
Qy	1419	GATCCATGAGCTTAACAGCGCCAAAGTGAACAGCATGATGATACCGTTGGCGGAGGCT	1478
Db	1406	GATCAAGAGCTTAAAGTGGCAATGTGACCAAGTCATGAGAGAGATTTGGTGGAACT	1465
Qy	1479	TGGGAGCTCCCTGGGCGCACTGTGGACCTGGGCTGACAGCAACCAATGAGTGTGCA	1538
Db	1466	TGGGATCTCTGGGCGCACTGTGACCTGGGCTTAACAGCGGATCTTGGTGTGCTCA	1525
Qy	1539	CCCCAGTCTTGGGAGAGTGGCCAAAGTTTGGAGGAGACACCTACTGCTTGGCGGAGC	1598
Db	1526	CCCCAGCTTGGGAGAGTGGCCAAAGCTTGAAGAAATGTCCATTTTGGCGGAGC	1585
Qy	1599	CGTGAGACACTGCTGCTCCGCTTTGGCAAGACCATCATGAGGAGCAGCTGGTACTGAA	1658
Db	1586	TGTTGAGACCTGTTACTCCGTTTGGAAAGAACATTTAGAGAACAGCTGGTGTAAA	1645
Qy	1659	GGGGTGGCCAAACATCTCATCAACTGTATGGCATGAGGCGCTGCTGTCGCGGGCCAG	1718
Db	1646	GGGGTAGCCAAAGCTCTCATCAACTGTATGGCATGCTGCTGCTGAGAGCCAG	1705
Qy	1719	CGCTCCATCCGATTTGGCTCCGCAAGCACGACGAGGTTCTTCTTGGCCAAACACTT	1778
Db	1706	CGCTCCATCCGATTTGGCTCCGCAAGCACGATCATGATCAGAGGTTCTTGGCCAAACACTT	1765
Qy	1779	CTGCTGGAAGCTTACTTTCAGAAATCTCTTTCAGCCTCTCTAGCTGAGCAAGTATGCTCC	1838
Db	1766	CTGCTGGAAGCTTATTTCCAGAAATCTCTTTCAGCCTCTCTAGCTGAGCAAGAGCTCC	1825
Qy	1839	AGAAACCTTAGATGAGCAGATTAAGAAAGTTCAGCAGAGATCCTTGAAGAGCGAGCCTA	1898
Db	1826	AGAAACCTTAGATGAGCAGATTAAGAAAGTTCAGCAGAGATCCTTGAAGAGCGAGCCTA	1885
Qy	1899	TATCTGTGCCACCTCTTGACAGGACATGCTGAGCAGGGGACA	1943
Db	1886	TATCTGTGCCACCTCTTGACAGGACATCCTGAGGCTGGAACCA	1930
RESULT 4			
AK028571			
LOCUS		2906 bp	linear
DEFINITION		Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:4732402K02 product:VERY-LONG-CHAIN ACYL-COA	
ACCESSION		AK028571.1	GI:26080919
VERSION		AK028571.1	GI:26080919
KEYWORDS		HTC; CAP trapper	
SOURCE		Mus musculus (house mouse)	
ORGANISM		Mus musculus	
REFERENCE			
AUTHORS		Carninci, P. and Hayashizaki, Y.	
TITLE		High-efficiency full-length cDNA cloning	
JOURNAL		Meth. Enzymol. 303, 19-44 (1999)	
MEDLINE		99279253	
PUBMED		10349636	
REFERENCE			
AUTHORS		Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	
TITLE		Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
JOURNAL		Genome Res. 10 (10), 1617-1630 (2000)	
MEDLINE		20499374	
PUBMED		11042159	
REFERENCE			
AUTHORS		Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	
TITLE		RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer	
JOURNAL		Genome Res. 10 (11), 1757-1771 (2000)	
MEDLINE		20530913	
PUBMED		11076861	
REFERENCE			
AUTHORS		Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamakawa, I., Saito, T., Okazaki, Y., Gijobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gliszi, C., King, B., Kochiwa, H.,	

QY	1593	CGGACCGTGGAGACACTGCTCGCTTGGCAAGACCATCATGAGGAGCAGCTGGT	1652
Db	2362	CGGACTGTGAGACCTGTTACTCGTTTGGAAAGAACATGTAGAGAACACTGGT	2421
QY	1653	ACTGAAGCGGTGGCCAAACATCCTCATCAACCTGTATGGCATGACGGCGTGTCTGTCGG	1712
Db	2422	GTAAAGCGGTAGCCAAACATCCTCATCACTTGTATGGCATGACTGTCTGTCGAG	2481
QY	1713	GGCCAGCGGTCCATCCGATTTGGGCTCGGCAACACGACGAGGTCTCTTGGCCAA	1772
Db	2482	AGCCAGCGGTCCATCAGAAATTTGGGCTCAGGAATCATGATCAGAGGTCTCTTGGCCAA	2541
QY	1773	CACCTTCTCGTGGAGGCTTACTTTCAGAAATCTCTTCAGCCCTCTCTCAGCTGGACAAATTA	1832
Db	2542	CATGTTCTGTGTGGAGGCTTATTTCCAGAAATCTCTTCAGCCCTCTCTCAGCTGGACAAATTA	2601
QY	1833	TGCTCCAGAAACCTAGATGAGCAGATTAAAGAAAGTGTCCAGCAGATCCTTTCAGAGCG	1892
Db	2602	CGCTCCAGAAACCTGGATGAGCAGATTAAAGAAAGTGTCTCGGCAGATCCTTTCAGAGCG	2661
QY	1893	AGCTATATCTGTGCCACCCCTGTGACAGGACATCTGAGCGAGGACA	1943
Db	2662	AGCTATATCTGTGCCACCCCTGTGACAGGACATCTGAGCGGCTGGAACCA	2712
RESULT 5			
LOCUS	AF078854	1123 bp	linear
DEFINITION	Homo sapiens NPD002 mRNA, complete cds.		
ACCESSION	AF078854		
VERSION	AF078854.1 GI:5531822		
KEYWORDS	HTC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1123) Hu, R., Peng, Y., Song, H., Huang, Q., Mao, Y., Zhang, Q., Mao, M., Fu, G., Luo, M. and Chen, J.		
TITLE	Human NPD002 mRNA, complete cds		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1123)		
AUTHORS	Hu, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-JUL-1998) Shanghai Institute of Endocrinology, Rui-Jin Hospital, Molecular Medical Center, 197 Rui-Jin Road II,, Shanghai 200025, P. R. China		
FEATURES	Location/Qualifiers		
source	1..1123		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/tissue_type="normal pituitary"		
	303..821		
	/codon_start=1		
	/product="NPD002"		
	/protein_id="AAD44486.1"		
	/db_xref="GI:5531823"		
	/translation="MEQLVLKPGGQHPQVWHDGRAVAGQPLHPWAPQPRPGSL GOHLRSLAESLQPLSAGQVCSKRPSRLRKCPKPSRLRSEPISVPTLMTGHAEAR GQCPLLPAPHGFLDDCYFFRRLILGLSQKPFVPLHLKGLAWESLFOVLTOR QCSLTGPSQLLN"		
BASE COUNT	270 a	297 c	303 g 253 t
ORIGIN			
Query Match	41.8%	Score 1026;	DB 11; Length 1123;
Best Local Similarity	98.3%;	Pred. No. 1.9e-214;	
Matches 1101;	Conservative 0;	Mismatches 10;	Indels 9; Gaps 6;
QY	1337	TCTTCGAGGAAACCAATGAGATTCTCCGGATGTACATCGCCCTGACGGGTCTGCAGCATG	1396
Db	9	TCTTCGAGGAAACCAATGAGATTCTCCGGATGTACATCGCCCTGACGGGTCTGCAGCATG	68

QY	1397	OCGCGCCGATCCTGACTACCAAGGATCCATGAGCTTAACAGGCCAACAGTGGACAGCTCA	1456
Db	69	-CGCGCCGATCCTGACTACCAAGGATCCATGAGCTTAACAGGCCAACAGTGGACAGCTCA	127
QY	1457	TGGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCGGAACCTGTGGACCTGGGGCTGA	1516
Db	128	TGGATACCGTTGGCGGCGG---TCAGGGAAGTCCCTGGGCGGAACCTGTGGACCTGGGGCTGA	183
QY	1517	CAGCAACCATGGAGTTGTGCACCCCACTCTTGGCGGAGAGTGCACAAAGTTTGAGGAGA	1576
Db	184	CAGCAACCATGGAGTTGTGCACCCCACTCTTGGCGGAGAGTGCACAAAGTTTGAGGAGA	243
QY	1577	ACACCTACTGCTTGGCGCGGAGGCTGGAGACACTGCTCGCTCCGCTTTTGGCAAGACCATCA	1636
Db	244	ACACCTACTGCTTGGCGCGGAGGCTGGAGACACTGCTCGCTTTTGGCAAGACCATCA	303
QY	1637	TGGAGGACAGCTGTACTGAAAG-CGGGTGGCCAACTCTCTCACTCACTGATGGCATG	1695
Db	304	TGGAGGACAGCTGTACTGAAAGCGGGTGGCCAACTCTCTCACTCACTGATGGCATG	363
QY	1696	ACGGCCGTGCTGTCGGCGGAGCGCTCCATCCGCAATTTGGCTCCGCAACACGACACAC	1755
Db	364	ACGGCCGTGCTGTCGGCGGAGCGCTCCATCCGCAATTTGGCTCCGCAACACGACACAC	423
QY	1756	GAGGTTCTTGGCCAAACACCTTCTCGGTGGAGGCTTACTTGCAGAACTCTTTCAGCCTC	1815
Db	424	GAGGTTCTTGGCCAAACACCTTCTCGGTGGAGGCTTACTTGCAGAACTCTTTCAGCCTC	483
QY	1816	TCTCAGCTGCAGAACTGATGCTCCAGAAACCTTAGAT-GAGCAGATTAAAGAAAGTGTCCCA	1874
Db	484	TCTCAGCTGCAGAACTGATGCTCCAGAAACCTTAGATCGAGCAGATTAAAGAAAGTGTCCCA	543
QY	1875	GCAGATCCTTGAGAGCGAGCGCTATATCTGTGCCCACTCTTGGCAGACATGCTGAGG	1934
Db	544	GCAGATCCTTGAGAGCGAGCGCTATATCTGTGCCCACTCTTGGCAGACATGCTGAGG	603
QY	1935	C-AGGGGACAGTGTCCCTGCTACGCCCGCCCTACCCATGCGCCGTTTGTCTGGATGACT	1993
Db	604	CTAGGGGACAGTGTCCCTGCTACGCCCGCCCTACCCATGCGCCGTTTGTCTGGATGACT	663
QY	1994	GTTACTCTTTTTCAGAAAGG-TGTTGGGATTATACAGGTTAAGCTTTTGTCTCCCGTC	2052
Db	664	GTTACTCTTTTTCAGAAAGGTTTGTGGGATTATACAGGTTAAGCTTTTGTCTCCCGTC	723
QY	2053	TGCACCTGAAGGTTGCTCGCTGGCGTGGAGAGCGCTCTTCCAGGTTTTCAGCTGCAGGC	2112
Db	724	TGCACCTGAAGGTTGCTCGCTGGCGTGGAGAGCGCTCTTCCAGGTTTTCAGCTGCAGGC	783
QY	2113	AGTGTCTCTTAACAGGACCATCACAGCTTCTGAACTGAGCCGGAGAGAGATGGAATT	2172
Db	784	AGTGTCTCTTAACAGGACCATCACAGCTTCTGAACTGAGCCGGAGAGAGATGGAATT	843
QY	2173	GCTGACCCCTGGAACCTGGCGGTTATTTCTGGTCAATTGAGGAGACACCATAGTGGAACTGG	2232
Db	844	GCTGACCCCTGGAACCTGGCGGTTATTTCTGGTCAATTGAGGAGACACCATAGTGGAACTGG	903
QY	2233	GGCTTATGCTGCTCCAGGTTGAGGTTGGGTGGGACCTGTGTCAAGTGTGGATAG	2292
Db	904	GGCTTATGCTGCTCCAGGTTGAGGTTGGGTGGGACCTGTGTCAAGTGTGGATAG	963
QY	2293	CCATTTCTGCTCAACACCATTTCTCTAAGAAACAGCTTGAAGCTCTGTCTGGGTCAAT	2352
Db	964	CCATTTCTGCTCAACACCATTTCTCTAAGAAACAGCTTGAAGCTCTGTCTGGGTCAAT	1023
QY	2353	CATTTAACTAGAGAGGAGGCACTTAAACATGTACAGGAACCATTTTAAACAAAGATA	2412
Db	1024	CATTTAACTAGAGAGGAGGCACTTAAACATGTACAGGAACCATTTTATACAAAGATA	1083
QY	2413	TAAATGTCACAACTCTGTGTACTGTTTAAAAAATAAAAAA	2452
Db	1084	TAAATGTCACAACTCTGTGTACTGTTTAAAAAATAAAAAA	1123

Invitrogen. This sequence belongs to sequence cluster 2225.f For more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DF034AA06Qp1&cluster=2225.f>. Contact : Feng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Paraday Avenue Genoscope sequence ID : CS0DF034AA06Qp1.

FEATURES		Location/Qualifiers	
source	1..1201	/organism="Homo sapiens"	
		/mol_type="mrna"	
		/db_xref="taxon:9606"	
		/clone="CS0DF034Y11"	
		/tissue_type="FETAL BRAIN"	
		/dev_stage="fetal"	
		/clone_lib="Homo sapiens FETAL BRAIN"	
		/note="Organ: brain; Vector: pcMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pcMVSPORT 6 vector. Library was not normalized."	
BASE COUNT		320 a	266 c 306 g 244 t 65 others
ORIGIN			
Query Match		38.4%	Score 940.4; DB 9; Length 1201;
Best Local Similarity		94.7%	Pred. No. 1.2e-195;
Matches		983;	Conservative 27; Mismatches 23; Indels 5; Gaps 4;
QY	32	GAGACTGAGGTGAGGTGGGGAACATCGGCGCAGCATGAGCGCTGGCGG-GCTCTTCCG 90	
DB	63	GGGATTAGGCTGAGGCTGGGGAACATCGGCGCAGCATGAGCGGTGGGNGCTTCTCTG 122	
QY	91	CGCACCGGCTGGGCTCGGCTCGCGGCTGCTGTCTCTACCGCGGAACCGCGG 150	
DB	123	CGCACCGGCTGCGCTGCTGCTGCGGCGTCTGTGTCTCTACCGCGGAACCGCGG 182	
QY	151	CTACTGCCACCGCGGCTGTAGAGCTTTCGCCAAAGAGCTTTCTTAGGCAATC 210	
DB	183	CTACTGCCACCGCGGCTGTAGAGCTTTCGCCAAAGAGCTTTCTTAGGCAATC 242	
QY	211	RAGAGAAGAGTTTCCCAATTTCCAGAGTTAGCCAAAGTAACTTAATGAATCAAT 270	
DB	243	RAGAGAAGAGTTTCCCAATTTCCAGAGTTAGCCAAAGTAACTTAATGAATCAAT 302	
QY	271	CAGTTCTTGGACCGGTGGAAAAATCTTCTACTGAAGGTGGACTCCGGAATAATGAC 330	
DB	303	CAGTTCTTGGACCGGTGGAAAAATCTTCTACTGAAGGTGGACTCCGGAATAATGAC 362	
QY	331	CAGGAAGGAATCCAGACTTGGAGAAATTTGAGAGCTAGGCTTTTGGG 390	
DB	363	CAGGAAGGAATCCAGACTTGGAGAAATTTGAGAGCTAGGCTTTTGGG 422	
QY	391	CTGCAAGTCCCAAGAAATATGGTGGCTGGGCTTCTCCACACACATGACTCAAGACTA 450	
DB	423	CTGCAAGTCCCAAGAAATATGGTGGCTGGGCTTCTCCACACACATGACTCAAGACTA 482	
QY	451	GGGAGATCATCAGATGGTGGTCCATCTGACCTGGGAGCGCCAGCGCTATT 510	
DB	483	GGGAGATCATCAGATGGTGGTCCATCTGACCTGGGAGCGCCAGCGCTATT 542	
QY	511	GGCTCAAGGGATCATCTTGGCTGGCACTGAGGAGCAAGAAAGCCAAATCTTGCTTAA 570	
DB	543	GGCTCAAGGGATCATCTTGGCTGGCACTGAGGAGCAAGAAAGCCAAATCTTGCTTAA 602	
QY	571	CTGGCTCGGGGAGACATTCGACCTTCTCCCTCAGCGCCAGCGAGTGGGCGAT 630	
DB	603	CTGGCTCGGGGAGACATTCGACCTTCTCCCTCAGCGCCAGCGAGTGGGCGAT 662	
QY	631	GCAGCTCAATCCGAGCAGACACACTAAGTGAAGCAAGAACACTTACATCTCAAT 690	
DB	663	GCAGCTCAATCCGAGCAGACACACTAAGTGAAGCAAGAACACTTACATCTCAAT 722	
QY	691	GGTCCAAAGTCTGGATTACTAATGGAGGACTGGCCAAATATTTTACTGTGTTGCAAG 750	

DB	723	GGCTCCAAAGTCTGGATTACTTAATGGAGGACTGGCCCAATATTTTACTGTGTTGCAAG 782	
QY	751	ACTGAGCTCGTGTGATCTGTGATGATCAAGCAACAAATCAAGCAATTCATAGTAA 810	
DB	783	ACTGAGCTCGTGTGATCTGTGATGATCAAGCAACAAATCAAGCAATTCATAGTAA 842	
QY	811	AGAGACTTTGGTGGAGTCACTAATGGGAAACCCCAAGATAAATAGGATTCGGGGCTCC 870	
DB	843	AGAGACTTTGGTGGAGTCACTAATGGGAAACCCCAAGATAAATAGGATTCGGGGCTCC 902	
QY	871	AACACTTGTGAGTCCATTTTCAAAACACCAAGATACCTGTGAAACATCCTTTGGAGAG 930	
DB	903	AASASTTGTGAAGTCSATTTTGAACACACSAAGATACSTGTGGAAACATCCTTTGAGAG 962	
QY	931	GTCGAGATGGGTTTAAGTGGCCATGAACATCCTCAACAGCGCGGCTTCAGCATGGGC 990	
DB	963	GTCGAGATGGGTTTAAGTGGCCATGAACATCCTCAACAGCGCGGCTTCAGCATGGGC 1022	
QY	991	AGCGTCTGGTGGCTCTCAAGAGATGATGAAATGACTGTGAGTACGCTGCACACA 1050	
DB	1023	ASC-KCGTGGCTGGGCTG-TYAARAGATTATTTGWAATTTT--YTGAAGTACCTGCCCA 1078	
QY	1051	AGSAAACAGTTTAAACAAG 1068	
DB	1079	AGGAAGTTTANAARAG 1096	

RESULT 8
LOCUS BX458691 Homo sapiens PLACENTA Homo sapiens linear EST 22-MAY-2003
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION BX458691
VERSION BX458691.1 GI:31025045
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2225.f For
more information about this cluster, see
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DE007AF120P1&cluster=2225.f>. Contact :
Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0DE007AF120P1.

FEATURES		Location/Qualifiers	
source	1..1201	/organism="Homo sapiens"	
		/mol_type="mrna"	
		/db_xref="taxon:9606"	
		/clone="CS0DE007YK23"	
		/tissue_type="PLACENTA"	
		/clone_lib="Homo sapiens PLACENTA"	
		/note="Vector: pcMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the pcMVSPORT 6 vector. Library was not normalized."	
BASE COUNT		290 a	284 c 337 g 240 t 50 others
ORIGIN			
Query Match		38.2%	Score 937.8; DB 13; Length 1201;

Best Local Similarity 97.6%; Pred. No. 4.4e-195; Matches 983; Conservative 9; Mismatches 11; Indels 4; Gaps 4;									
QY	778	GTGAAGACAAAATACACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGG	137						
Db	66	GTGAAGACAAAATACACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGG	125						
QY	838	AAACCCGAGATAAATAGGCATTCGGGGCTCCCAACACTTGTGAAGTCCATTTCGAAA	897						
Db	126	AAACCCGAGATAAATAGGCATTCGGGGCTCCCAACACTTGTGAAGTCCATTTCGAAA	185						
QY	898	ACCAAGATACCTGTGGAACATCCTTGGAGAGTGGAGATGGGTTTAAAGTGGCCAT	957						
Db	186	ACCAAGATACCTGTGGAACATCCTTGGAGAGTGGAGATGGGTTTAAAGTGGCCAT	245						
QY	958	AACATCCTCAACAGCGCGCGTTCAGCATGGCAGCGTGGTGGCTGCTCAAGACA	1017						
Db	246	AACATCCTCAACAGCGCGSGTTCAGCATGGCAGCGTGGTGGCTGCTCAAGACA	305						
QY	1018	TTGATTGAATGACTGCTGAGTACGGCTGCACAGAAACAGTTTAAAGAGGCTCAAT	1077						
Db	306	TTGATTGAATGACTGCTGAGTACGGCTGCACAGAAACAGTTTAAAGAGGCTCAAT	365						
QY	1078	GAATTTGGATGATTACAGAGAAATTTGCACATGATGGCTCAGAGGCTTACGTATG	1137						
Db	366	GAATTTGGATGATTACAGAGAAATTTGCACATGATGGCTCAGAGGCTTACGTATG	425						
QY	1138	AGTATGACCTACTCACAGCAGGATGCTGGACCAACCTGGCTTCCCGACTGTCTCATC	1197						
Db	426	AGTATGACCTACTCACAGCAGGATGCTGGACCAACCTGGCTTCCCGACTGTCTCATC	485						
QY	1198	GAGGAGCCATGTTGAAGGTGTTACAGTCCGAGCGCGCTGGCAGTGTGAGTGAGCG	1257						
Db	486	GAGGAGCCATGTTGAAGGTGTTACAGTCCGAGCGCGCTGGCAGTGTGAGTGAGCG	545						
QY	1258	CTCAGATCCTCGGGGGCTTGGGCTACACAGGGACTATCCGTACGACCGCATACTGCGT	1317						
Db	546	CTCAGATCCTCGGGGCTTGGGCTACACAGGGACTATCCGTACGACCGCATACTGCGT	605						
QY	1318	GACACCCGATCTCTCATCTTCAGGGAACCAATGATATTCCTCGGATGATACATGCC	1377						
Db	606	GACACCCGATCTCTCATCTTCAGGGAACCAATGATATTCCTCGGATGATACATGCC	665						
QY	1378	CTGACGGGTCTGCAATATGC-CGGCGCGATCTGACTACCAAGGATCCATGAGCTTAAACA	1436						
Db	666	CTGACGGGTCTGCAATATGC-CGGCGCGATCTGACTACCAAGGATCCATGAGCTTAAACA	725						
QY	1437	GGCCAAAGTGAACAGTATGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGCGCG	1496						
Db	726	GGCCAAAGTGAACAGTATGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGCGCG	784						
QY	1497	AACTGTGGACCTGGGGCTGACAGCAACCATGGAGTTGTGCACCCAGTCTTGGGACAG	1556						
Db	785	AACTGTGGACCTGGGGCTGACAGCAACCATGGAGTTGTGCACCCAGTCTTGGGACAG	844						
QY	1557	TGCCAACAGTTTGAAGAACACCTACTGCTTCGGCCGACCGTGGAGACACTGTGCT	1616						
Db	845	TGCCAACAGTTTGAAGAACACCTACTGCTTCGGCCGACCGTGGAGACACTGTGCT	904						
QY	1617	CGCTTTGGCAAGACCATCATGGAGGACAGCTGGTACTGAAGCGGTGGCCAAATCCT	1676						
Db	905	CGCTTTGGCAAGACCATCATGGAGGACAGCTGGTACTGAAGCGGTGGCCAAATCCT	964						
QY	1677	CATCAACCTGATGGCATGAGGGCGTGTGTCGGGGCCAGCGCTCCATCCGCATTGG	1736						
Db	965	CATCAACCTGATGGCATGAGGGCGTGTGTCGGGGCCAGCGCTCCATCCGCATTGG	1024						
QY	1737	GCTCCGAACACACAGGAGTTCCTTTGGCCAAACACCTCTTGGC 1783							
Db	1025	GST-CGCAACMACGA-CACGAGGKTTTYYTGGCCAAACACTTTTSGT 1069							
RESULT 9									

AL567735	1201 bp	mrna	linear	EST 31-MAY-2003	
LOCUS					
DEFINITION	AL567735 Homo sapiens FETAL BRAIN Homo sapiens CDNA clone				
ACCESSION	CS0DF034YALL 3-PRIME, mRNA sequence.				
VERSION	AL567735				
KEYWORDS	AL567735.2 GI:31290607				
SOURCE	EST.				
ORGANISM	Homo sapiens (human)				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1 (bases 1 to 1201)				
JOURNAL	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
COMMENT	Full-length cDNA libraries and normalization Unpublished On Feb 16, 2001 this sequence version replaced gi:12921390. Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr was not normalized. Library was constructed by Life Technologies, a division of invitrogen. This sequence belongs to sequence cluster 2225.f For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0DF034AA06NP1&cluster=2225.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paradise Avenue Genoscope sequence ID : CS0DF034AA06NP1. Location/Qualifiers 1. 1201 /organism="Homo sapiens" /mol_type="mrna" /db_xref="taxon:9606" /clone="CS0DF034YALL" /tissue_type="FETAL BRAIN" /dev_stage="fetal" /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized." BASE COUNT 271 a 302 c 295 g 270 t 63 others ORIGIN Query Match 35.9%; Score 879.2; DB 9; Length 1201; Best Local Similarity 96.5%; Pred. No. 3.2e-182; Matches 953; Conservative 17; Mismatches 11; Indels 7; Gaps 7; QY 1316 GTGACACCCGATCCTCTCTATCTTCGAGGGGAACCAATGAGATTCCTCGGATGTACATCG 1375 Db 983 SKGAMACCCGATCCTCTCTAT-YTCGAGGGAACCAATGAGATTCCTCGGATGTACATCC 925 QY 1376 CCTGACGGGTCTGC-AGCATCCGCGCGCATCTCTGACTACCAAGATCCATGAGCTT-AA 1433 Db 924 GCCTGACGGGTCTGCAAGCATGCCGCGCGCATCTCTGACTACCAAGATCCATGAGCTTAAA 865 QY 1434 ACAGGCCAAAGTGAGCAGACATGATACCGTTGGCCGAGGCTTCGGGACTCCCTGGG 1493 Db 864 ACAGGCCAAAGTGAGCAGACATGATACCGTTGGCCGAGGCTTCGGGACTCCCTGGG 805 QY 1494 CCGAACTGTGGACCTGGGGCTGACAGGCAACCATGGAGTTGTGACCCAGCTCTTGGGGA 1553 Db 804 CCGAACTGTGGACCTGGGGCTGACAGGCAACCATGGAGTTGTGACCCAGCTCTTGGGGA 745 QY 1554 CAGTGCCCAACAGTTTGAAGGAGAACACCTACTGCTTCGGCCGAGCCGTTGGAGACACTGCT 1613 Db 744 CAGTGCCCAACAGTTTGAAGGAGAACACCTACTGCTTCGGCCGAGCCGTTGGAGACACTGCT 685 QY 1614 GCTCCGCTTTGGCAAGACCATCATGGAGGACAGCTGTGTAAGCGGTTGGCCAAACAT 1673 Db 684 -STCCGCTTTGGCAAGACCATCATGGAGGACAGCTGTGTAAGCGGTTGGCCAAACAT 626				

```
Qy 1674 CCTCATCAACCTGTATGGCATGACGGCCGCTGCTGCGGGCCAGCGCTCCATCCGAT 1733
Db 625 CCTCATCAACCTGTATGGCATGACGGCCGT-STGTGCGGGCCAGCGCTCCATCCGAT 567
Qy 1734 TGGGCTCCGCAACCCAGCACAGGCTTCTTTGGCCACACCTTCTGCTGGTGAAGCTTA 1793
Db 566 T-GGTCGGCAACCCAGCACAGGCTTCTTTGGCCACACCTTCTGCTGGTGAAGCTTA 508
Qy 1794 CTTGCAAGATCTCTTCAGGCTCTCTCAGCTGGACAAAGTATGCTCCAGAAAACCTAGATGA 1853
Db 507 CTTGCAAGATCTCTTCAGGCTCTCTCAGCTGGACAAAGTATGCTCCAGAAAACCTAGATGA 448
Qy 1854 CGAGATTAGAAAGTGTCCAGCAGATCTTTGAGAAAGCAGCTTATCTGTGCCACCC 1913
Db 447 CGAGATTAGAAAGTGTCCAGCAGATCTTTGAGAAAGCAGCTTATCTGTGCCACCC 388
Qy 1914 TCTGGACAGACATGCTGAGGAGGAGGACAGTGTCCCTGCTACCGCCGCCCTTACCAC 1973
Db 387 TDTGGACAGACATGCTGAGGAGGAGGACAGTGTCCCTG-TACCGCCGCCCTTACCAC 329
Qy 1974 TGGCCCGTGTGCTGGATGACTGTACTCTTTTTCAGAAAGTGTGGGATTATCACAGTT 2033
Db 328 TGGCCCGTGTGCTGGATGACTGTACTCTTTTTCAGAAAGTGTGGGATTATCACAGTT 269
Qy 2034 AAGCCTTTTGTCTCCCGTCTGCACCTGAAGGTTGTGCGCTGGCGTGGAGAGCCTCTTC 2093
Db 268 AAGCCTTTTGTCTCCCGTCTGCACCTGAAGGTTGTGCGCTGGCGTGGAGAGCCTCTTC 209
Qy 2094 CAGGTTTTCAGCTGAGGAGTGTCTCTTAACAGGACCATCACAGTTCTTGAAGTACCC 2153
Db 208 CAGGTTTTCAGCTGAGGAGTGTCTCTTAACAGGACCATCACAGTTCTTGAAGTACCC 149
Qy 2154 GGAGAGAGAGATGGAATGCTGACCCCTGGAACCTGGCGGTATCTGCTCATTCAGGAG 2213
Db 148 GGAGAGAGAGATGGAATGCTGACCCCTGGAACCTGGCGGTATCTGCTCATTCAGGAG 89
Qy 2214 ACACCATAGTGGAACTGGGCTTATGCTGCTGCCCTCCAGGCTGTGAGTGGTGGGAC 2273
Db 88 ACACCATAGTGGAACTGGGCTTATGCTGCTGCCCTCCAGGCTGTGAGTGGTGGGKCC 29
Qy 2274 CTGCTGTCAGTGTGGATACCCATTTCTG 2301
Db 28 KTGTGTCWGGTGTGGATAKKTTTTRK 1

RESULT 10
AL549834
LOCUS
DEFINITION
AL549834 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1054Y017 5-PRIME, mRNA sequence.
ACCESSION
AL549834
VERSION
AL549834.2 GI:31271652
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 969)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12886203.
Contact: Genoscope
Genoscope, Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2225.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1054AH09Qp1&cluster=2225.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
```

Faraday Avenue Genoscope sequence ID : CS0D1054AH09Qp1.

FEATURES
source

```
1..969
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="CS0D1054Y017"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

BASE COUNT
ORIGIN

274 a 215 c 269 g 203 t

Query Match 35.5%; Score 869.6; DB 9; Length 969;
Best Local Similarity 98.6%; Pred. No. 4e-180;

Matches 891; Conservative 6; Mismatches 5; Indels 2; Gaps 2;

Qy 15 TGCAGCGCTAAGAGAGGAGAGACTGAGGCTGAGGCTGGGGAACATCGGCGAGCATGAGCGG 74

Db 67 TGCAGAGCTAAGAGAGGAGGAGACTGAGGCTGAGGCTGGGGAACATCGGCGAGCATGAGCGG 126

Qy 75 CTGCGG-GCTCTTCCTCGGCACACAGGCTCGGCTCGCTCGCGGGGCTTGGTGGTCT 133

Db 127 CTGCGGNGCTTCTCTCGGCACACAGGCTCGGCTCGCTCGCGGGGCTTGGTGGTCT 186

Qy 134 CTACCGCGAAACCGCGGCTACTGCGACACAGCCGCGCTGTACGAGCTTTCGCCAAAGAGC 193

Db 187 CTACMGCGAACCGGAGGATACTGCGCACACAGCCGCGCTGTACGAGCTTTCGCCAAAGAGC 246

Qy 194 TTTTCTAGGCAAAATCAAGAAAGAAAGTATTTTCCATTTCAGAAAGTTAGCAAGATG 253

Db 247 TTTTCTAGGCAAAATCAAGAAAGAAAGTATTTTCCATTTCAGAAAGTTAGCAAGATG 306

Qy 254 AACTTAATGAATCAATCAGTCTTGGGACCCGCTGGAAAATTTCTTCACTGAAGAGGTGG 313

Db 307 AACTTAATGAATCAATCAGTCTTGGGACCCGCTGGAAAATTTCTTCACTGAAGAGGTGG 366

Qy 314 ACTCCCGAAAAATTGACCGAGGAGGAAATCCAGATGAAATTTGGAGAAATTTGAAGA 373

Db 367 ACTCCCGAAAAATTGACCGAGGAGGAAATCCAGATGAAATTTGGAGAAATTTGAAGA 426

Qy 374 GCCTAGGCGCTTTTGGGCTGCAAGTCCAGAGAATATGTCGCTGGCTTCTCCAACA 433

Db 427 GCCTAGGCGCTTTTGGGCTGCAAGTCCAGAGAATATGTCGCTGGCTTCTCCAACA 486

Qy 434 CCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCACTGTGACCCCTGG 493

Db 487 CCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCACTGTGACCCCTGG 546

Qy 494 CAGCGCACCGAGCTATTGGCCTCAAGGGATCATCTTGGCTGGCACTGAGGAGCAGAAG 553

Db 547 CAGCGCACCGAGCTATTGGCCTCAAGGGATCATCTTGGCTGGCACTGAGGAGCAGAAG 606

Qy 554 CCAAACTACTTGCCTAAACTGGCTCCGGGAGCACATTCAGCCTTCTGCCCTCAAGGAGC 613

Db 607 CYAAATACTTGCCTAAACTGGCTCCGGGAGCACATTCAGCCTTCTGCCCTCAAGGAGC 666

Qy 614 CAGCCAGTGGGAGCGATGAGCCTCAATCCGAGAGAGCCACACTAAGTGAAGACAGA 673

Db 667 CAGCCAGTGGGAGCGATGAGCCTCAATCCGAGAGAGCCACACTAAGTGAAGACAGA 726

Qy 674 AGCATTACATCTCAATGGCTCAAGGCTGTGATTACTAATGGAGACTGGGCAATATT 733

Db 727 AGCATTACATCTCAATGGCTCAAGGCTGTGATTACTAATGGAGACTGGGCAATATT 786

Qy 734 TTACTGTGTTTGCAGAGACTGAGGTGCTTGTGATTCTGTGATGATGAGGAGGAGGAGTCA 793

Db 787 TTACTGTGTTTGCAGAGACTGAGGTGCTTGTGATTCTGTGATGATGAGGAGGAGGAGTCA 846

Qy 794 CAGCATTCATAGTGAAGAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAGATAAT 853

```

Db      847 CAGCATTCATAGTAAAGAGACTTTGGTGGAGTCACTAATGGGAACCCCAAGATAAAT 906
Qy      854 TAGGCATTCGGGGCTCCAACTTGTGAAGTCCATTTTGAACACCAAGATACCTGTGG 913
Db      907 TAGGCATTCGGGGCT-CMACACTTGTGAAGTCCATTTTGAACACCAAGATACCTGTGG 965
Qy      914 AAAA 917
Db      966 AAAA 969

RESULT 11
BM552756
LOCUS      1010 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_6573876 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5477224
5', mRNA sequence.
ACCESSION BM552756
VERSION   BM552756.1 GI:18790885
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1010)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: ATCC
          CDNA Library Preparation: Rubin Laboratory
          DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLCMI994 row: o column: 17
          High quality sequence stop: 711.
          Location/Qualifiers
            1..1010
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:5477224"
              /tissue_type="astrocytoma grade IV, cell line"
              /lab_host="DH10B (phage-resistant)"
              /clone_lib="NIH_MGC_98"
              /note="Organ: brain; Vector: pOTB7; Site:1: XhoI; Site:2:
              ECORI; cDNA made by oligo-dT priming. Directionally
              cloned into EcoRI/XhoI sites using the following 5'
              adaptor: GGCAGGAG(G). Library constructed by Ling Hong
              in the laboratory of Gerald M. Rubin (University of
              California, Berkeley) using ZAP-cDNA synthesis kit
              (Stratagene) and Superscript II RT (Life Technologies).
              Note: this is a NIH_MGC Library."
          229 a 289 c 285 g 206 t 1 others

BASE COUNT 229 a 289 c 285 g 206 t 1 others
ORIGIN

Query Match
Best Local Similarity 35.4%; Score 867.2; DB 12; Length 1010;
Matches 909; Conservative 0; Mismatches 34; Indels 2; Gaps 2;

Qy      1042 CCCTGCACAAGCAAGCTTTAAACAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAA 1101
Db      1 GCCTGCACAAGCAAGCTTTAAACAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAA 60
Qy      1102 TTTGCATGATGGCTCAGAAGGCTTACCTCATGGAGAGATGACCTACCTACACAGCAGGG 1161
Db      61 TTTGCATGATGGCTCAGAAGGCTTACCTCATGGAGAGATGACCTACCTACACAGCAGGG 120
Qy      1162 ATGCTGCACCAACCTGGCTTCCCGACTGCTCCATCGAGCAGCCATGGTGAAGGTTC 1221

```

```

Db      121 ATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGACCCATGCTGAAGGTGTTTC 180
Qy      1222 AGCTCCGAGGCGCCCTGGCAGTGTGTGAGTGAGGGCTTCAGATCCTCGGGGGCTTGGGC 1281
Db      181 AGCTCCGAGGCGCCCTGGCAGTGTGTGAGTGAGGGCTTCAGATCCTCGGGGGCTTGGGC 240
Qy      1282 TACACAAGGAGCTATCCGTACGAGGCGCATACCTGCGTGACACCCGCGATCCTCCTCATCTTC 1341
Db      241 TACACAAGGAGCTATCCGTACGAGGCGCATACCTGCGTGACACCCGCGATCCTCCTCATCTTC 300
Qy      1342 GAGGGAACCAATGAGATTCCTCCGGATGTACATCGCCCTGACGGGTCTCGACGATCCCGGC 1401
Db      301 GAGGGAACCAATGAGATTCCTCCGGATGTACATCGCCCTGACGGGTCTCGACGATCCCGGC 360
Qy      1402 CGCATCCTGACTACCAAGATCCAGCTTAAACAGGCGCAAGTGCAGACAGTCAATGAT 1461
Db      361 CGCATCCTGACTACCAAGATCCAGCTTAAACAGGCGCAAGTGCAGACAGTCAATGAT 420
Qy      1462 ACCGTTGGCGGAGGCTTCCGGACTCCCTGGGCGCAACTGTGGACCTGGGCTGACAGGC 1521
Db      421 ACCGTTGGCGGAGGCTTCCGGACTCCCTGGGCGCAACTGTGGACCTGGGCTGACAGGC 480
Qy      1522 AACCATGAGTGTGCACCCAGTCTTCGGGACAGTGCACAAGTTTGGAGGAAACACC 1581
Db      481 AACCATGAGTGTGCACCCAGTCTTCGGGACAGTGCACAAGTTTGGAGGAAACACC 540
Qy      1582 TACTGCTTTCGGCGGAGCGGTGGAGACACTGCTGCTCGCTTTGGCAAGACCATCATCGAG 1641
Db      541 TACTGCTTTCGGCGGAGCGGTGGAGACACTGCTGCTCGCTTTGGCAAGACCATCATCGAG 600
Qy      1642 GAGCAGCTGGTACTGAAGCGGTGGCAACATCTCATCAACCTGTATGGCATGACGCGC 1701
Db      601 GAGCAGCTGGTACTGAAGCGGTGGCAACATCTCATCAACCTGTATGGCATGACGCGC 660
Qy      1702 GTGCTGTCCGGCGGCGCAGCGCTCCATCGGATTCGGCTCCGCAACACAGTGTGGAGGAGTT 1761
Db      661 GTGCTGTCCGGCGGCGCAGCGCTCCATCGGATTCGGCTCCGCAACACAGTGTGGAGGAGTT 720
Qy      1762 CTCTTGGCAACACCTTCTCGGTGGAGTCTTACTTTCAGAGTCTCTTTCAGCCCTCTCTCAG 1821
Db      721 CTCTTGGCGCCACACCTTCTCGGTGGAGTCTTACTTTCAGAGTCTCTTTCAGCCCTCTCTCAG 780
Qy      1822 CTGGACAAGTATGCTCCAGAAAACCTAGATGAGCAGATTGAAGAAGTGTCCAGCAGAT - 1880
Db      781 CTGGACAAGTATGCTCCAGAAAACCTAGATGAGCAGATTGAAGAAGTGTCCAGCAGATC 840
Qy      1881 CCTTGAAGCGAGGCTTATATCTGTG-CCACACCTCTTGGACAGGACATGCTGAGGCAGGG 1939
Db      841 CCTGGAGAGCAGGCTTATATCTGTGCCCCCACCCTCTTGGACAGGACATGCTTGAACCAAG 900
Qy      1940 GACAGTGTCCCTCTGCTACCGCCCGCCCTACCCATGCCCGCTTGC 1984
Db      901 GGCACATGTCCCTCTGCTACCGCCCGCCCTTACCCCTTACCCCATGGC 945

RESULT 12
BM561207
LOCUS      1079 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_6589858 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5475994
5', mRNA sequence.
ACCESSION BM561207
VERSION   BM561207.1 GI:18806273
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1079)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-r@mail.nih.gov

```

Tissue Procurement: DCTD/DP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
High quality sequence stop: 700.
Location/Qualifiers
1. .1079
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5475994"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_41"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."
BASE COUNT 295 a 245 c 311 g 226 t 2 others
ORIGIN

Query Match 35.0%; Score 859.4; DB 12; Length 1079;
Best Local Similarity 96.0%; Pred. No. 6.9e-178;
Matches 924; Conservative 0; Mismatches 32; Indels 6; Gaps 4;
QY 53 GAACATCGGCAGCATGAGCGCTCGGGCTCTCTCGGCACACCGGCTGCGGCTCGTG 112
DB 1 GAACATCGGCAGCATGAGCGCTCGGGCTCTCTCTCGGCACACCGGCTGCGGCTCGTG 60
QY 113 CCTCGCGGGTCTGTGTCTCTACCGCGAACCGGGCTTACTGCGCACCGCGCGCTG 172
DB 61 CCTCGCGGGTCTGTGTCTCTACCGCGAACCGGGCTTACTGCGCACCGCGCGCTG 120
QY 173 TACGAGCTTTCGCAAGAGCTTTCTTAGGCAAAATCAAGAAAGAGTTTCCAT 232
DB 121 TACGAGCTTTCGCAAGAGCTTTCTTAGGCAAAATCAAGAAAGAGTTTCCAT 180
QY 233 TTCAGAAAGTTAGCAAGATCAACTTAATGAATCAATCAGTTCTTGGGACCCCGTGAAA 292
DB 181 TTCAGAAAGTTAGCAAGATCAACTTAATGAATCAATCAGTTCTTGGGACCCCGTGAAA 240
QY 293 AATTCTTCATGAAGAGTGGACTCCCGAAATTTGACGAGGAGGAAATCCAGATG 352
DB 241 AATTCTTCATGAAGAGTGGACTCCCGAAATTTGACGAGGAGGAAATCCAGATG 300
QY 353 AAACCTTCGAGAAATGAAGAGCTTAGGGCTTTTGGGCTCAAGTCCCAAGAAATATG 412
DB 301 AAACCTTCGAGAAATGAAGAGCTTAGGGCTTTTGGGCTCAAGTCCCAAGAAATATG 360
QY 413 GTGGCTGGGCTTCTCCAAACCATGTACTCAAGCATAGGGGAGATCATCAGATGATG 472
DB 361 GTGGCTGGGCTTCTCCAAACCATGTACTCAAGCATAGGGGAGATCATCAGATGATG 420
QY 473 GTTCATCATGTGACCTTGGCAGCGCACCGGCTATTGGCTCAAGGGGATCATCTGG 532
DB 421 GTTCATCATGTGACCTTGGCAGCGCACCGGCTATTGGCTCAAGGGGATCATCTGG 480
QY 533 CTGGCACTGAGGAGCAGAAAGCAATATCTTGCCTAAATCTGCCTCCGGGAGCACATG 592
DB 481 CTGGCACTGAGGAGCAGAAAGCAATATCTTGCCTAAATCTGCCTCCGGGAGCACATG 540
QY 593 CAGCCTTCGCTCAGGAGCAGCCAGTGGGAGGATGCAGCTCATCTCCGAGCAG 652
DB 541 CAGCCTTCGCTCAGGAGCAGCCAGTGGGAGGATGCAGCTCAATCCGAGCAG 600

QY 653 CCACATAAGTGAAGACAAGACACTACATCTCTCAATGGCTCAAGGTCTGATTA 712
DB 601 CCACATAAGTGAAGACAAGACACTACATCTCTCAATGGCTCAAGGTCTGATTA 660
QY 713 ATGAGAGCTGGCCCAATATTTTACTCTGTTGCAAGACTGAGTCTGTTGATTCGATG 772
DB 661 ATGAGAGCTGGCCCAATATTTTACTCTGTTTGGAAAGACTGAGTCTGTTGATTCGATG 720
QY 773 GATCAGTGAAGACAAGAAATCACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTA 832
DB 721 GATCAGTGAAGACAAGAAATCACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTA 780
QY 833 ATGGAAACCCGAAGATTAATAGGCATTC -GGGGCTCCACACACTTGTGAA -GTCCATTT 890
DB 781 ATGGAAACCCGAAGATTAATAGGCATTCGGGGGCTCCACACACTTGTGAAAGTCCATTT 840
QY 891 TGAAGAACCAAGATACCTGTGAAACACATCTTGG -AGAGTCCGAGATGGGTTAA - 947
DB 841 TGAAGAACCAAGATACCTGTGAAACACATCTTGGGAAACATCTCGGANATGGTTTAAAG 900
QY 948 -GGTGGCCATGAACATCTCAACAGCGCGGTTTCAGCATGGGAGCTCTGGTGGCGC 1006
DB 901 GTGGCCATGGAACATCTCTCAAGCGGGCGGTTTAAAGCATGGGAGCGGCTCTGGCG 960
QY 1007 TG 1008
DB 961 TG 962

RESULT 13
AL582366/c
LOCUS
DEFINITION
AL582366 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
Homo sapiens cDNA clone CS0DL006YK09 3-PRIME, mRNA sequence.
AL582366
VERSION
AL582366.2 GI:31320583
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1206)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 16, 2001 this sequence version replaced gi:12950278.
COMMENT
Contact: Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2225.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DL006AF05NP1cluster-2225.f. Contact :
Feng Liang Email : fliang@lifetech.com URL : Invitrogen Corporation 1600
http://fulllength.invitrogen.com/ Parady Avenue Genoscope sequence ID : CS0DL006AF05NP1.
Location/Qualifiers
1. .1206
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL006YK09"
/cell_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(GT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 276 a 306 c 292 g 259 t 73 others

FEATURES
source

ORIGIN

Query Match 34.8%; Score 854; DB 9; Length 1206;
 Best Local Similarity 96.9%; Pred. No. 1.le-176;
 Matches 901; Conservative 10; Mismatches 15; Indels 4; Gaps 4;

QY 1345 GGAACCAATGATTTCTCCGATGTACATCCCTCGAGGCTCTGACGATCGCGGCCG 1404
 DB 930 GGAACAAGARATTTCTCGGARTGTACATYCSCTAMGGGTCTCAASAATACCGT-CGC 872

QY 1405 ATCTGTACTACACGATCCATGCTTAACAGG-CCAAAGTGAAGACATCATGGATAC 1463
 DB 871 ATCTGTACTACCMGAYCCATGAGCTTAACAGGCCCAAGTGAAGACATCATGGATAC 812

QY 1464 CGTTGGCGGAGGTTTGGGACTCCCTGGCGGACATGTGAGCTGGGGTGCACAGCAA 1523
 DB 811 CGTTGGCGGAGGTTTGGGACTCCCTGGCGGACATGTGAGCTGGGGTGCACAGCAA 752

QY 1524 CCATGGAGTTGTGCACCCAGTCTTGGGACAGTGCACCAAGTTTGGAGAACACCTA 1583
 DB 751 CCATGGAGTTGTGCACCCAGTCTTGGGACAGTGCACCAAGTTTGGAGAACACCTA 692

QY 1584 CTGCTTGGCGGACCTGTGAGACACTGTCTGCTCCGCTTGGCAAGACCATCATGAGGA 1643
 DB 691 CTGCTTGGCGGACCTGTGAGACACTGTCTGCTCCGCTTGGCAAGACCATCATGAGGA 632

QY 1644 GCAGCTGGTACTGAAGCGGTGGCCAAACATCTCATCAACCTGTATGGCATGACGCCGT 1703
 DB 631 GCAGCTGGTACTGAAGCGGTGGCCAAACATCTCATCAACCTGTATGGCATGACGCCGT 572

QY 1704 GCTGTCCGGCGGACGCTTCCATCCGATTTGGCTCCGCAACCGACGACGAGTTCT 1763
 DB 571 G-TGTCGGCGGACGCTTCCATCCGATTTGG-TCCGCAACCGACGACGAGTTCT 514

QY 1764 CTGTCGCAACACTTCTGCTGGAAGCTTACTTGGAGATCTCTTCAGCTCTCTCAGCT 1823
 DB 513 CTGTCGCAACACTTCTGCTGGAAGCTTACTTGGAGATCTCTTCAGCTCTCTCAGCT 454

QY 1824 GGACAAGTATGCTCCAGAAACCTAGATGAGCAGATTAAGAAAGTGTCCAGCAGATCCT 1883
 DB 453 GGACAAGTATGCTCCAGAAACCTAGATGAGCAGATTAAGAAAGTGTCCAGCAGATCCT 394

QY 1884 TGAGAAGCAGCTATATCTGTGCCACCTCTCGACAGGACATCTCGAGCGGGGACA 1943
 DB 393 TGAGAAGCAGCTATATCTGTGCCACCTCTCGACAGGACATCTCGAGCGGGGACA 334

QY 1944 GTGTCCCTGTACCGCGGCCCTTACCATGGCCCGTTGCTGGATGACTTTACTCTTT 2003
 DB 333 GTGTCCCTGTACCGCGGCCCTTACCATGGCCCGTTGCTGGATGACTTTACTCTTT 274

QY 2004 TTTCAGAGGTGTTGGGATTTATCAGGTTAAGCTTTTGTCCCGCTCTGCACCTGAAG 2063
 DB 273 TTTCAGAGGTGTTGGGATTTATCAGGTTAAGCTTTTGTCCCGCTCTGCACCTGAAG 214

QY 2064 GTTGTCCCTGTGCGGAGGCTCTTCCAGGTTTGTACCTCGAGCAGTGTCTCTA 2123
 DB 213 GTTGTCCCTGTGCGGAGGCTCTTCCAGGTTTGTACCTCGAGCAGTGTCTCTA 154

QY 2124 ACAGACCATCACAGCTTCTCACTGAGCGCGGAGAGAGATGGAATGCTGACCCCTG 2183
 DB 153 ACAGACCATCACAGCTTCTCACTGAGCGCGGAGAGAGATGGAATGCTGACCCCTG 94

QY 2184 GAACCTGGGGGTATTTGGTCAITTAGAGAGACACCATAGTGGAAACTGGGGCTTATGCTG 2243
 DB 93 GAACCTGGGGGTATTTGGTCAITTAGAGAGACACCATAGTGGAAACTGGGGCTTATGCTG 34

QY 2244 CTGCTCCAGGCTGTGAGTGGGTGGGAC 2273
 DB 33 CTGCTCCAGGCTGTGAGTGGGTGGGAC 4

RESULT 14
 BQ938039

LOCUS BQ938039 890 bp mRNA linear EST 21-AUG-2002
 DEFINITION AGENCOURT_894875 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6469922
 5', mRNA sequence.
 ACCESSION BQ938039
 VERSION BQ938039.1 GI:22353517
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 1 (bases 1 to 890)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 890)
 NIH-MGC http://mgc.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLNL1999 row: n column: 03
 High quality sequence stop: 611.
 Location/Qualifiers
 1. .890
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6469922"
 /tissue_type="leiomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_71"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo df.
 Average insert size 2.1 kb."
 BASE COUNT 242 a 214 c 248 g 186 t
 ORIGIN

Query Match 34.8%; Score 852.4; DB 13; Length 890;
 Best Local Similarity 99.6%; Pred. No. 2.3e-176;
 Matches 886; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

QY 13 CCTGCGCGCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGAACATCGGGCAGCATGAGC 72
 DB 1 CCTGCGCGCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGAACATCGGGCAGCATGAGC 60

QY 73 GGCTGGGGCTCTTCTTGGCGCACACCGCTGCGGCTCGCTGCGGGGCTGTGGTGGTC 132
 DB 61 GGCTGGGGCTCTTCTTGGCGCACACCGCTGCGGCTCGCTGCGGGGCTGTGGTGGTC 120

QY 133 TCTACCGCGAACCGGGGCTACTGCGCACCGCCGCTGTACGAGCTTTCGCCAAGAG 192
 DB 121 TCTACCGCGAACCGGGGCTACTGCGCACCGCCGCTGTACGAGCTTTCGCCAAGAG 180

QY 193 CTTTCTTAGGCAAAATCAAGAAAGAGTTCCTCCCATTTCCAGAGTTAGCCAAAGAT 252
 DB 181 CTTTCTTAGGCAAAATCAAGAAAGAGTTCCTCCCATTTCCAGAGTTAGCCAAAGAT 240

QY 253 GAACCTAATGAATCAATCAGTTCTTGGACCCGCTGGAAATTTCTTCACTGAAGAGCTG 312
 DB 241 GAACCTAATGAATCAATCAGTTCTTGGACCCGCTGGAAATTTCTTCACTGAAGAGCTG 300

QY 313 GACTCCGCAAAATTTGACCAAGGAGGAAATCCAGATGAACCTTTGGAGAAATTTGAAG 372
 DB 301 GACTCCGCAAAATTTGACCAAGGAGGAAATCCAGATGAACCTTTGGAGAAATTTGAAG 360

QY 373 AGCCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGGCTGGGCTTCTCCAAAC 432
 DB 361 AGCCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGGCTGGGCTTCTCCAAAC 420

QY 433 ACCATGTACTCAGACTAGGGGAGATCATCAGATGGATGGTCCATCATCTGTGACCCCTG 492

digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized.

BASE COUNT	329 a	261 c	309 g	245 t	57 others
ORIGIN					
Query Match	34.8%; Score 852.4; DB 9; Length 1201;				
Best Local	Similarity 96.1%; Pred. No. 2.4e-176;				
Matches	903; Conservative 12; Mismatches 18; Indels 7; Gaps 4;				
QY	11	TCCCTGCGGCGCTAAGAGGGGAGACTGAGCTGAGCTGGGAAACATCGGAGCATGA	70		
DB	60	TCCCTGCGGCGCTAAGAGGGGAACTGAGGCTGAGGCTGGGAAACATCGGAGCATGA	119		
QY	71	CGGCTGCGG-GCTCTTCTCGCCACACAGGCTCGGCTCGTGCCTGCCGGGTCTGGTG	129		
DB	120	CGGCTGCGGNGCTCTTCTCGCCACACAGGCTCGGCTCGTGCCTGCCGGGTCTGGTG	179		
QY	130	GTCTCTACCGGAACCGCGGCTACTGCGCACAGCGCGCTGTACGAGCTTTCGCCAAA	189		
DB	180	GTCTCTACCGGAACCGCGGCTACTGCGCACAGCGCGCTGTACGAGCTTTCGCCAAA	239		
QY	190	GAGCTTTTCCCTAGGCAAAATCAAGAAGAAAGTTCCTCCATTTCCAGAAAGTTAGCAA	249		
DB	240	RAGCTTTTCCCTAGGCAAAATCAAGAAGAAAGTTCCTCCATTTCCAGAAAGTTAGCAA	299		
QY	250	GATGAACCTTAATGAATCAATCAGTCTTGGGACCGCTGGAAATTTCTTCACTGAAGAG	309		
DB	300	GATGAACCTTAATGAATCAATCAGTCTTGGGACCGCTGGAAATTTCTTCACTGAAGAG	359		
QY	310	GTGGACTCCGCAAAATTTGACGAGGAAGGAAATCCGAGATGAAACTTTTGAGAAATTTG	369		
DB	360	GTGGACTCCGCAAAATTTGACGAGGAAGGAAATCCGAGATGAAACTTTTGAGAAATTTG	419		
QY	370	AAGAGCCTTAGGCTTTTGGGCTGCAAGTCCAGAAATATGTTGGCTGGCTTCCTCC	429		
DB	420	AAGAGCCTTAGGCTTTTGGGCTGCAAGTCCAGAAATATGTTGGCTGGCTTCCTCC	479		
QY	430	AACACCATCTACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCATCTGTGACC	489		
DB	480	AACACCATCTACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCATCTGTGACC	539		
QY	490	CTGGCAGCCACACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCTTCAGGAGAG	549		
DB	540	CTGGCAGCCACACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCTTCAGGAGAG	599		
QY	550	ARAGCCAAATATCTTGCTTAACTGGCTCCGGGAGACATTCGACGCTTCTGCCTCAGG	609		
DB	600	ARAGCCAAATATCTTGCTTAACTGGCTCCGGGAGACATTCGACGCTTCTGCCTCAGG	659		
QY	610	GAGCCAGCCAGTGGGAGCGATGAGCTCAATCCGGAGCAGACCATTAAGTGAAGAC	669		
DB	660	GAGCCAGCCAGTGGGAGCGATGAGCTCAATCCGGAGCAGACCATTAAGTGAAGAC	719		
QY	670	AAGAAGCCTACATCTCAATGCTCCAGGCTCGATTACTTAATGGAGCTGGCCAAAT	729		
DB	720	AAGAAGCCTACATCTCAATGCTCCAGGCTCGATTACTTAATGGAGCTGGCCAAAT	779		
QY	730	ATTTTACTCTGTTTGCACAAAGACTAGGCTGTTGATCTGTAGTGAAGTCAAGTGAAG	789		
DB	780	ATTTTACTCTGTTTGCACAAAGACTAGGCTGTTGATCTGTAGTGAAGTCAAGTGAAG	839		
QY	790	ATCACACATCTACATAGTGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAAGT	849		
DB	840	ATCACACATCTACATAGTGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAAGT	899		
QY	850	AAATTAGGCATTCGGGCTCCAAACACTTTGTGAAGTCCATTTTGAACACACAGATACCT	909		
DB	900	AAATTAGGCATTCGGGCTCCAAACACTTTGTGAAGTCCATTTTGAACACACAGATACCT	955		
QY	910	GTGGAACACATCTTGGAGAGGTCGGAGATGGGTTTAAAG	949		
DB	956	GTGGAACACATCTTGGAGAGGTCGGAGATGGGTTTAAAG	993		

Db	421	ACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCAGCTGTGACCCTG	480
QY	493	GCAGCCACAGGCTATTGGCTCAAGGGATCATCTTGGCTGGGACTGAGGAGCAGAAA	552
DB	481	GCAGCCACAGGCTATTGGCTCAAGGGATCATCTTGGCTGGGACTGAGGAGCAGAAA	540
QY	553	GCCAAATACTTGCCTAAACTGGCTCCGGGGAGCAGATTCGACGCTTCTGCTCAGGAG	612
DB	541	GCCAAATACTTGCCTAAACTGGCTCCGGGGAGCAGATTCGACGCTTCTGCTCAGGAG	600
QY	613	CCAG-CCAGTGGAGGATGAGCTCAATCCGGAGCAGACACACTAAGTGAAGACAA	671
DB	601	CCAGCCAGTGGAGGATGAGCTCAATCCGGAGCAGACACACTAAGTGAAGACAA	660
QY	672	GAAGCACTACATCTCAATGGGCTCCAGGCTCGGATTTACTAATGGAGGACTGGCCAAAT	731
DB	661	GAAGCACTACATCTCAATGGGCTCCAGGCTCGGATTTACTAATGGAGGACTGGCCAAAT	720
QY	732	TTTTACTGTTTGCAGAACTGAGTCTGTTGATCTGATGATGATGATGATGATGATGAT	791
DB	721	TTTTACTGTTTGCAGAACTGAGTCTGTTGATCTGATGATGATGATGATGATGATGAT	780
QY	792	CACAGATTCATAGTGAAGACACTTT-GGTGAGTCACTAATGGGAAACCCGAGATA	850
DB	781	CACAGATTCATAGTGAAGACACTTTGGTGGAGTCACTAATGGGAAACCCGAGATA	840
QY	851	AATTAGGCAATCGGGCT-CCAACTTTGGAAGTCACTTTGAAACAC	899
DB	841	AATTAGGCAATCGGGCTCCAACTTTGGAAGTCACTTTTGAACAC	890

RESULT 15
AL560956
LOCUS
DEFINITION
AL560956 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
Homo sapiens cDNA clone CS0DL006YK09 5-PRIME, mRNA sequence.
ACCESSION
AL560956
VERSION
AL560956.2 GI:31285085
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (bases 1 to 1201)
Li W.B., Gruber C., Jessee J. and Polayes D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12907913.
Contact: Genoscope
Genoscope - Centre National de Sequenage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2225.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DL006AF05QPI&cluster=2225.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0DL006AF05QPI.

FEATURES
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL006YK09"
/cell_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was

Tue Sep 9 10:14:37 2003

us-09-945-326-1.rst

Page 17

Search completed: September 7, 2003, 01:51:41
Job time : 5168.14 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 13:30:07 ; Search time 62 Seconds
(without alignments)
1375.251 Million cell updates/sec

Title: US-09-945-326-2
Perfect score: 3153
Sequence: 1 MSGCGFLFRTTAAARACRGL.....SQOILEKRAYICAHPLDRTC 621

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513375 seqs, 137303645 residues

Total number of hits satisfying chosen parameters: 513375

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2.6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2.6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2.6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2.6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2.6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2.6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2.6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2.6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2.6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2.6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2.6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2.6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2.6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2.6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2.6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2.6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2.6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2.6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3153	100.0	621	10	US-09-945-326-2
2	3153	100.0	621	15	US-10-168-274-24
3	659.5	20.9	646	15	US-10-156-761-10104
4	650	20.6	384	15	US-10-272-419-6
5	615.5	19.5	385	15	US-10-156-761-12560
6	604	19.2	432	12	US-10-354-358-28
7	574.5	18.2	390	15	US-10-156-761-14447
8	562.5	17.8	386	15	US-10-156-761-12809
9	556	17.6	383	15	US-10-156-761-14127
10	530.5	16.8	401	15	US-10-156-761-9454
11	495	15.7	385	15	US-10-156-761-12814
12	492.5	15.6	409	15	US-10-156-761-14657
13	480	15.2	513	12	US-10-181-319-33
14	479	15.2	415	15	US-10-205-823-4
15	452.5	14.4	376	12	US-10-238-075-951

16	395	12.5	346	10	US-09-925-300-1645	Sequence 1645, Ap
17	382.5	12.1	403	15	US-10-156-761-8886	Sequence 8886, Ap
18	374	11.9	375	15	US-10-156-761-9843	Sequence 9843, Ap
19	367.5	11.7	797	11	US-10-156-761-14069	Sequence 14069, A
20	340	10.8	797	11	US-09-999-314A-2	Sequence 2, Appl1
21	339	10.8	413	15	US-10-156-761-8259	Sequence 8259, Ap
22	330	10.5	407	15	US-10-156-761-14142	Sequence 14142, A
23	310	9.8	421	15	US-10-156-761-8920	Sequence 8920, Ap
24	298.5	9.5	601	9	US-09-815-242-5070	Sequence 5070, Ap
25	297.5	9.4	386	15	US-10-156-761-11399	Sequence 11399, A
26	295	9.4	156	11	US-09-999-314A-7	Sequence 7, Appl1
27	289	9.2	608	15	US-10-156-761-11924	Sequence 11924, A
28	283	9.0	398	10	US-09-738-626-3821	Sequence 3821, Ap
29	275	8.7	392	15	US-10-156-761-8806	Sequence 8806, Ap
30	260	8.2	393	15	US-10-156-761-14669	Sequence 14669, A
31	254.5	8.1	585	10	US-09-976-059-26	Sequence 26, Appl1
32	227.5	7.2	400	12	US-09-769-724-50	Sequence 50, Appl1
33	222	7.0	379	15	US-10-156-761-13723	Sequence 13723, A
34	207	6.6	384	15	US-10-156-761-11373	Sequence 11373, A
35	207	6.6	429	15	US-10-156-761-11746	Sequence 11746, A
36	204.5	6.5	376	15	US-10-156-761-14670	Sequence 14670, A
37	203	6.4	1261	15	US-10-156-761-11188	Sequence 11188, A
38	202	6.4	71	9	US-09-864-761-47228	Sequence 47228, A
39	199.5	6.3	549	15	US-10-156-761-9661	Sequence 9661, Ap
40	196.5	6.2	383	15	US-10-156-761-11368	Sequence 11368, A
41	190	6.0	413	15	US-10-156-761-11355	Sequence 11355, A
42	181.5	5.8	553	10	US-09-976-059-25	Sequence 25, Appl1
43	178.5	5.7	414	15	US-10-119-651-6	Sequence 6, Appl1
44	173	5.5	609	15	US-10-156-761-11187	Sequence 11187, A
45	169.5	5.4	394	15	US-10-156-761-8745	Sequence 8745, Ap

ALIGNMENTS

RESULT 1

US-09-945-326-2
; Sequence 2, Application US/09945326
; Patent No. US20020127880A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Hunter, John Joseph
; TITLE OF INVENTION: 62112, A NOVEL HUMAN DEHYDROGENASE AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: MNI-187
; CURRENT APPLICATION NUMBER: US/09/945, 326
; CURRENT FILING DATE: 2001-08-31
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 621
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-326-2

Query Match	100.0%	Score 3153	DB 10	Length 621
Best Local Similarity	100.0%	Pred. No. 3.4e-298		
Matches 621	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	MSGCGFLFRTTAAARACRGLVVTANRRLLRTPSPVAFKELFLGKIKKKEVPPFPVS	60	
DB	1	MSGCGFLFRTTAAARACRGLVVTANRRLLRTPSPVAFKELFLGKIKKKEVPPFPVS	60	
QY	61	QDELNEINQFGLPVEKFTTEEDSRKIDQEKIPDETLEKLSGLFLQVPEEYGLGF	120	
DB	61	QDELNEINQFGLPVEKFTTEEDSRKIDQEKIPDETLEKLSGLFLQVPEEYGLGF	120	
QY	121	SNMTYSRLGEIISMDGSITVTLLAAHQALGLKGIILAGTEEQAKYLPKLASEHIAFCL	180	
DB	121	SNMTYSRLGEIISMDGSITVTLLAAHQALGLKGIILAGTEEQAKYLPKLASEHIAFCL	180	

QY 181 TEPASGSDAASIRSRATLSEDKKHILNGSKVWITNGGLANIFTVFAKTEVDSGSKVD 240
DB 181 TEPASGSDAASIRSRATLSEDKKHILNGSKVWITNGGLANIFTVFAKTEVDSGSKVD 240
QY 241 KITAFIVERDRGGVTNGKPKDLGIRGNTCEVHFENTKIPVENILGEVGDGFKVAMNIL 300
DB 241 KITAFIVERDRGGVTNGKPKDLGIRGNTCEVHFENTKIPVENILGEVGDGFKVAMNIL 300
QY 301 NSGRFSMGSVVAGLLKRLIEMTAETACRQKFNKRLSEFGLIOEKFALMAQKAYVMESMT 360
DB 301 NSGRFSMGSVVAGLLKRLIEMTAETACRQKFNKRLSEFGLIOEKFALMAQKAYVMESMT 360
QY 361 YLTAGMLDQPGFPCDSEAAWVFSSEAAWQCVSEALQILGGLGYTRDYPYERILDRTR 420
DB 361 YLTAGMLDQPGFPCDSEAAWVFSSEAAWQCVSEALQILGGLGYTRDYPYERILDRTR 420
QY 421 ILLIFEGTNEILRMVIALTGLQHAGRIILTRIHLELKOAKVSTVMDTVGRRRLDSLGRTVD 480
DB 421 ILLIFEGTNEILRMVIALTGLQHAGRIILTRIHLELKOAKVSTVMDTVGRRRLDSLGRTVD 480
QY 481 LGLTGNHGVVHPSLADSANKFEENTYCFGRVETILLRFGKTIMEEQVLKRVANILNL 540
DB 481 LGLTGNHGVVHPSLADSANKFEENTYCFGRVETILLRFGKTIMEEQVLKRVANILNL 540
QY 541 YGMTAVLSRASRSIRIGLRNHDHVEVLLANTFCVEAYLQNLFLSOLDKYAPENLDEQIKK 600
DB 541 YGMTAVLSRASRSIRIGLRNHDHVEVLLANTFCVEAYLQNLFLSOLDKYAPENLDEQIKK 600
QY 601 VSQILEKRAYICAHPLDRTC 621
DB 601 VSQILEKRAYICAHPLDRTC 621
RESULT 2
US-10-168-274-24
; Sequence 24, Application US/10168274
; Publication No. US20030124106A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: HILLMAN, Jennifer
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LU, Dyung Aina M.
; TITLE OF INVENTION: HUMAN OXIDOREDUCTASE PROTEINS
; FILE REFERENCE: PF-0754 PCT
; CURRENT APPLICATION NUMBER: US/10/168, 274
; PRIOR FILING DATE: 2002-08-26
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 24
; LENGTH: 621
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030124106A1 5540437CD1
US-10-168-274-24
Query Match 100.0%; Score 3153; DB 15; Length 621;
Best Local Similarity 100.0%; Pred. No. 3.4e-298;
Matches 621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSGCGLFRTTAAARACGLVSTANRRLLRTSPVRAFAKELFLGKIKKKEVPPPEVS 60
DB 1 MSGCGLFRTTAAARACGLVSTANRRLLRTSPVRAFAKELFLGKIKKKEVPPPEVS 60
QY 61 ODELNEINQFLGPVEKFFTEEDSRKIDQEGKIPDETLEKLSGLGLQVPEEYGGGLG 120

DB 61 QDELNEINQFLGPVEKFFTEEDSRKIDQEGKIPDETLEKLSGLGLQVPEEYGGGLG 120
QY 121 SNTMYSRLEIISMDGSIITVTLAAHQAIKGLIILAGTEEQAKYLPKILASGEHIAAFCL 180
DB 121 SNTMYSRLEIISMDGSIITVTLAAHQAIKGLIILAGTEEQAKYLPKILASGEHIAAFCL 180
QY 181 TEPASGSDAASIRSRATLSEDKKHILNGSKVWITNGGLANIFTVFAKTEVDSGSKVD 240
DB 181 TEPASGSDAASIRSRATLSEDKKHILNGSKVWITNGGLANIFTVFAKTEVDSGSKVD 240
QY 241 KITAFIVERDRGGVTNGKPKDLGIRGNTCEVHFENTKIPVENILGEVGDGFKVAMNIL 300
DB 241 KITAFIVERDRGGVTNGKPKDLGIRGNTCEVHFENTKIPVENILGEVGDGFKVAMNIL 300
QY 301 NSGRFSMGSVVAGLLKRLIEMTAETACRQKFNKRLSEFGLIOEKFALMAQKAYVMESMT 360
DB 301 NSGRFSMGSVVAGLLKRLIEMTAETACRQKFNKRLSEFGLIOEKFALMAQKAYVMESMT 360
QY 361 YLTAGMLDQPGFPCDSEAAWVFSSEAAWQCVSEALQILGGLGYTRDYPYERILDRTR 420
DB 361 YLTAGMLDQPGFPCDSEAAWVFSSEAAWQCVSEALQILGGLGYTRDYPYERILDRTR 420
QY 421 ILLIFEGTNEILRMVIALTGLQHAGRIILTRIHLELKOAKVSTVMDTVGRRRLDSLGRTVD 480
DB 421 ILLIFEGTNEILRMVIALTGLQHAGRIILTRIHLELKOAKVSTVMDTVGRRRLDSLGRTVD 480
QY 481 LGLTGNHGVVHPSLADSANKFEENTYCFGRVETILLRFGKTIMEEQVLKRVANILNL 540
DB 481 LGLTGNHGVVHPSLADSANKFEENTYCFGRVETILLRFGKTIMEEQVLKRVANILNL 540
QY 541 YGMTAVLSRASRSIRIGLRNHDHVEVLLANTFCVEAYLQNLFLSOLDKYAPENLDEQIKK 600
DB 541 YGMTAVLSRASRSIRIGLRNHDHVEVLLANTFCVEAYLQNLFLSOLDKYAPENLDEQIKK 600
QY 601 VSQILEKRAYICAHPLDRTC 621
DB 601 VSQILEKRAYICAHPLDRTC 621
RESULT 3
US-10-156-761-10104
; Sequence 10104, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156, 761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10104
; LENGTH: 646
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10104
Query Match 20.9%; Score 659.5; DB 15; Length 646;
Best Local Similarity 30.6%; Pred. No. 4.3e-55;
Matches 186; Conservative 107; Mismatches 242; Indels 73; Gaps 16;
QY 38 AFKAFELPLGKIKKKEVPPPEVSQDELNELNLOFLGPVEKFFTEEDSRKIDQEGKIPDET 97
DB 34 SFAKELFLGRFLDLHPMPADEAQORGEFELAKLDFCETEDSARLEAREARIPDET 93


```

Db      266 YAKGVQERKQFGKAIADFQIQMLADMAKMEARQLTYAAAKSER-GDKDLTFQGA 324
QY      381 MVKVFSEAAQCVCSEALQILGGIGYRDYDYPYERILRDRTRILLIFEGTNEILRMVIA 437
Db      325 AAKCFASDVAMEVTTDAVOLLLGGYGYTRDYPVERMRDRAKITQIYEGTNQVQRIWVA 381

RESULT 6
US-10-354-358-28
; Sequence 28, Application US/10354358
; Publication No. US20030157082A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Hunter, John Joseph
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Lesoon, Andrea
; APPLICANT: Lightcap, Eric S.
; APPLICANT: Williamson, Mark
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
; TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
; TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
; TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
; TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
; TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
; TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 56428 MOLECULES
; FILE REFERENCE: MPI02-020PIRNMNIM
; CURRENT APPLICATION NUMBER: US/10/354,358
; CURRENT FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US 60/353,600
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/364,517
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/371,075
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/371,507
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/372,984
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/374,194
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/382,995
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/385,023
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/388,853
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/389,395
; PRIOR FILING DATE: 2002-06-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-354-358-28

Query Match      19.2%; Score 604; DB 12; Length 432;
Best Local Similarity 35.9%; Pred. No. 5.7e-50;
Matches 157; Conservative 81; Mismatches 155; Indels 44; Gaps 14;

QY      19 GLVSTANRRLRTPPPVAFKELFLGKIKKKEVFP-FPVSODE--LNEINQ--FLGP 73
Db      3 GLAV-----RLRGSELLR----RNFLTCLSSWKIPPHVSKSSQSEALLNITNNGIHFP 53
QY      74 VEKFTTEVDSRK-----IDEGKIPDETLEKLSGLGLGLOVPEYGG 117
Db      54 LOTFTDEEMMIKSVKKAQEOIAPLVSTMDENSKWEKSVIOGLFQOGLMGIEVDPEVGG 113

```

```

QY      118 LG--FSNTMYSRGELIISMDGSITVTLAAHQAIKGLKGIILAGTBEQKAKYLPKLASGEHI 175
Db      114 TGSFSLTVLV-IEELAKVDASVAVFCIEQNTLNTLIRKKGTEBEQKATYLPQUTT-EKV 171
QY      176 AAFCLTEPASGSDAASIRSRTLSEDKKHYYILNGSKVWITNGGLANIFTFVFAKTEVVDSD 235
Db      172 GSFCLSEAGAGSDSFALKTRADKEGD--YYVLNGSKMWISSAEHAGLFLVMANV-----D 224
QY      236 GSVKDK-ITAFIVERDFGVTNGKPEDKLGIRGNSNTCEVHFENTKIPVENILGEVGDGFK 294
Db      225 PTIGYKGTISFLVDROTGLGHGKPKENKLGRLRASSTCLPTFENVKVPANILGQIGHGYK 284
QY      295 VAMNINLSRFGSMGVVAGLLRLLEMTAEYACTRKQENKRLSEFGLIOEKFALMAQAKY 354
Db      285 YAIGSLNEGRIGTAAQMLGLAQGCCFYDTPYIKERIQFGKRLDFQGLQHOHVAVATQLE 344
QY      355 VMESMTYLTAGMLDQGFDPDCSTEAAMVKVFSSEAAWQCVSEALQILGLGLGTRDYPYER 414
Db      345 AARLLTYNAARLL-EAGKPFIK-EASMAKYVASEIAGQTTSCIEWMGGVGYTKDYPVEK 402
QY      415 ILRDRILLIFEGTNEI 431
Db      403 YFRDAKIGTIVEGASNI 419

RESULT 7
US-10-156-761-14447
; Sequence 14447, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14447
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14447

Query Match      18.2%; Score 574.5; DB 15; Length 390;
Best Local Similarity 34.4%; Pred. No. 3.6e-47;
Matches 135; Conservative 80; Mismatches 152; Indels 25; Gaps 8;

QY      55 PPEVVSQDELNEINQFLGPVEKFFTEEV--DSRKIDQEGKIPDETLEKLSGLGFLGLOVP 112
Db      6 POPVDRQLPTDEARDLISLRVIAQREIAPEAAEGEDAGHPREFVGLLSGLGLPYD 65
QY      113 EEVGGIGFNTMYSRGELIISMDGSITV-----TLAAHQAIKGLKGIILAGTBEQKAKY 165
Db      66 SEYGGGDQPYEYVLQALELAA-ARLTGVLGVSVHTLACH-----ALAQYGTKEQOQVEH 118
QY      166 LPKLASGEHIAFCLTEPASGSDAASIRSRTLSEDKKHYYILNGSKVWITNGGLANIFTV 225
Db      119 LPAMLGGGLGAYCLSEPSGSDAASLRKAV--RDGDDWVITGKAWITHGCIADFTV 176
QY      226 FAKTEVVDSDGSKVKDKITAFIVERDPPGVNTNGKPEDKLGIRGNSNTCEVHFENTKIPVNI 285
Db      177 MAAT-----GGEGARGITAFVPGDAEGLHAALPEKMKMGKSPTAQVHFQVGRVSDERR 231
QY      286 LGEVGDGFKVAMNINLSRFGSMGVVAGLLKRLLEMTAEYACTRKQENKRLSEFGLIOEK 345

```

```

Db      232  IGDEGQFAIALSALDSGRGLAACAVGVAQAALDEAVATATGQQGRIADPQGLRFEM 291
QY      346  FALMAKAYVNESYTLTACHLQDPGPPDCSIEAAMVKVFSSEAAQCSEALQILGGLG 405
Db      292  LADMATQIAGRAL-YLAAARLRDAGR-ESKQAAMAKLLCTDAAMKVTTTDAVQVLGGYG 349
QY      406  YTRDPYPERILDRTRILLIFEGTNEILRMVIA 437
Db      350  YTADFPVRYMREAKVLIQVEGTNQIQRMVIA 381

RESULT 8
US-10-156-761-12809
; Sequence 12809, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12809
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12809
Query Match      17.8%; Score 562.5; DB 15; Length 386;
Best Local Similarity 36.5%; Pred. No. 5.2e-46;
Matches 142; Conservative 72; Mismatches 144; Indels 31; Gaps 10;
QY      63  ELNEINQFLGPVEKFFTEVDSSRKID---QEGKIPDETLEKLSGLGLGLOVPEYGG 119
Db      8  ELEELRR---TVEE-FADVVPAPKIGDFYERHETPYEIVRMGRMLGFLPFPPEYGGM 63
QY      120 FSNMYSRLG-----EISMDSGITVTAAHQAIKGLGIIAGTEEQAKYLPKLAGSEHI 175
Db      64  GD---YALGIALLEELARVDSVAITLEAGVSLGAMPILHLFGTDAQAKAELPRLCSG 120
QY      176  AAFCLTEPASGSDAASIRSRATLSDEKHKHYILNGSKVWITNGG--LANIETVFAKTEVD 233
Db      121  GAFGLTEPDGSDAGATRTTARLDESTNEWVINGTKCFITNSGDTITGLTV---TAVTG 177
QY      234  SDGSKVKITAFIVERDFGVTNGKPEDKLGIRGNTCEVHFENTKIPVENILGEYCDGDF 293
Db      178  RKPDKPLISSIIVPSTGPTGTVAAAPYSKVGWNASDTRLSFADVRVPAANLGEQGR 237
QY      294  KVAMNINSGRFSMGVAVGLLRLIEMTAETAYACTRKFQFNKRLSEFGLIOEKFPALMAQ 353
Db      238  AQFLRLIDEGRIALATATLAQCQVDSVKYAGERHAFGRNIGAYQAIQFKIADMENKA 297
QY      354  YV-----MESMTYLTAGMLDQPPDCSIEAAMVKVFSSEAAQCSEALQILGGLGYTR 408
Db      298  HMAVGVGRDAASRLVAG---EP----FKKEAATKLYSSTVAVDNAREATQIHGGYGMN 350
QY      409  DYPVERTLRDRTRILLIFEGTNEILRMVIA 437
Db      351  EYPARMWRDSKLTLEIGECTSEVQRMVIA 379

RESULT 9

```

```

US-10-156-761-14127
; Sequence 14127, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14127
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14127
Query Match      17.6%; Score 556; DB 15; Length 383;
Best Local Similarity 34.6%; Pred. No. 2.2e-45;
Matches 134; Conservative 84; Mismatches 151; Indels 18; Gaps 9;
QY      63  ELNEINQFLGPVEKFFTE-EVDSRKI--DOEGKIPDETLEKLSGLGLGLOVPEYGG 119
Db      4  ELSEEQTAVALQALAKDFVDREAPNVIAWDRAEEVDSRLVKLGEVGLGLTVDEEYGG 63
QY      120 FSNMYSRLGEEIISM-DGSITVTAAHQAIKGLGIIAGTEEQAKYLPKLAGSEHIAAF 178
Db      64  GDHLAYCLVTEELGRCDSVSRGIVSVSLGVAKTIAYWGESEKQKRWLPGLTSGAVTGC 123
QY      179  CLTEPASGSDAASIRSRATLSDEKHKHYILNGSKVWITNGLANIFTVFAKTEVDSG 238
Db      124  GLTEPOTGSDAGALLTRAV--RDGDDYVYNGTKMETINGTWADVVLLEARS---TDAP 177
QY      239  KDKITAFIVERDFGVTNGKPEDKLGIRGNTCEVHFENTKIPVENILGEYCDGDFK 298
Db      178  HKGVSAFLVPTDTPGLTRPIHGKGLRCQATAEVLVEDVRVPASAMLAPEKGFSVAM 237
QY      299  ILNSGRFSMGVAVGLLRLIEMTAETAYACTRKFQFNKRLSEFGLIOEKFPALMAQ 358
Db      238  ALAKGRMSVAACGVGIAQAALDVAVKYATEREQFGKTTIAHQVLQVDELISDIADV 297
QY      359  MTLTAGMLD--QPGFPDCSIEAAMVKVFSSEAAQCSEALQILGGLGYTRDYPYERIL 416
Db      298  LTRVADLVDRQCP-FATASSQA---KLFASEAAVRAANNALQVPGGYGYDEYPA 353
QY      417  RTRILLIFEGTNEILRMVYI--ALTGL 441
Db      354  RDARVMTLYEGTSQIKLLIGRALTVG 380

RESULT 10
US-10-156-761-9454
; Sequence 9454, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262

```

Tue Sep 9 10:14:37 2003

Query Match 15.7%; Score 495; DB 15; Length 385;
Best Local Similarity 34.8%; Pred. No. 2e-39;
Matches 126; Conservative 63; Mismatches 141; Indels 32; Gaps 8;
CURRENT APPLICATION NUMBER: US/10/156.761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 9454
LENGTH: 401
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-9454

Query Match 16.8%; Score 530.5; DB 15; Length 401;
Best Local Similarity 34.8%; Pred. No. 7.3e-43;
Matches 136; Conservative 76; Mismatches 150; Indels 29; Gaps 11;
QY 64 LNEINQ-FLGPVEKEFTTEV--DSRKIDQEGKIPDETLEKLSLGLFLGLOVPEEYGGIGF 120
DB 10 LTDIOQELSTLVRDFVDKEIIPVATELEHREDEYPOOIYVDGKLGELGLGLMIPEEYGGIGE 69
QY 121 SNTMYSRLEGIISMDG-STIVTAAHQAIGLKGIIA-----GTEOKAKYLPKLASGE 173
DB 70 SLLTVALCEEIARGMWSVGIINTH-----FIVAYMLKOHGTQEQKHFLPRMALGE 122
QY 174 HIAAFCLTPEASGSDAASIRSRATLSSEKHHKILNGSKVWITNGLANIETVFAKTEVVD 233
DB 123 VRGAFMSPEALGSDVSATSSKAV--KDGEVLNGQKMLNGSSNLVAVLVRSDGSH 180
QY 234 SDGSVKDK ITAFIVERD--FG-----GVNKGXPKEDKLGIRGNTCEVHFENKIPVENIL 286
DB 181 PEGTAPHKSWTFLVEKEFGFGEVREGLTIPGKIDKMGYKGVDTTELIMDGLRVPANRVL 240
QY 287 -GEVGDGFKVAMNLSGRFSMGSVVAGLLKRLIETAYACTRQENKRLSEFGLIOFK 345
DB 241 GGTGTRGFYQMDGVEGRVNAARGCGVAQRAFLGVSYAQRHTFGKPIHQHQAIOFK 300
QY 346 FALMAOKAYVMESMTYLTAGMLDQPGFPDPCSTEAAMKVFSSFAAQCVSEALQILGGLG 405
DB 301 LAEMATKVEAHAMVNAARKKDSGERND--LEAGMAKYLASEYKEVVEDAFRIHGGYG 358
QY 406 YTRDYPYERILDRTRILLIFEGTNEILRMVYI 436
DB 359 FSKYEYERLYREAPMLLIGEGTAETQKMI 389

RESULT 11
US-10-156-761-12814
Sequence 12814, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156.761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 12814
LENGTH: 385
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-12814

Query Match 15.7%; Score 495; DB 15; Length 385;
Best Local Similarity 34.8%; Pred. No. 2e-39;
Matches 126; Conservative 63; Mismatches 141; Indels 32; Gaps 8;
QY 91 GRIPDETLEKLSLGLFLGLOVPEEYGGIG-----FSNTMY-----SRLEGIISMDGSITVTL 142
DB 36 GOAPRDFYKLAELGVFGIRVDEEHGAGIDYKKEAVLYETARAG--ISFGGS-----88
QY 143 AAHQATGLKGIILAGTEOKAKYLPKLASGEHIAAFCLTEPASGSDAASIRSRATLSSEK 202
DB 89 GVHVLGLPYIKLLATDEOKKRLPKFVSSEEMWAIAMTEPGTSDLAGMKTTAKLSDDG 148
QY 203 KHYLNGSKVWITNGLANIETVFAKTEVVDGSGVKDKITAFIVERDGGVTNGKPEDK 262
DB 149 THVVLNGAKTFTITGTVHADRMIVCARTSAPTKEDEK--RHGISLVFVDTKAEGISVSGKLDK 207
QY 263 LGTRGNTCEVHFENKIPVENILGEGVDFKVFAMNLSGRFSMGSVVAGLLKRLIEMT 322
DB 208 LGKTSDTALAFVDVKVPVEDLLGEEKNGFYLYLGHNLASERWGIAFGAYAKAAVREA 267
QY 323 AEYACTRKQPNKRLSEFGLIOEKPALM-----AOKAYVMESMTYLTAGMLDQPGFPDC 375
DB 268 KQYVQERVVEGQPVAAAFQ--NTRFELAAACQAEVDAEAADVDRATEALDAGEL-----T 318
QY 376 SIEAAMKVFSSFAAQCVSEALQILGGLGYTRDYPYERILDRTRILLIFEGTNEILRMV 435
DB 319 PAEASAKLFPATEVAHRVIDRCLQHGCGYGFMEYPIARLYADNRVNRIVGGTSEIMKTI 378
QY 436 IA 437
DB 379 IA 380

RESULT 12
US-10-156-761-14657
Sequence 14657, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156.761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 14657
LENGTH: 409
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-14657

Query Match 15.6%; Score 492.5; DB 15; Length 409;
Best Local Similarity 32.8%; Pred. No. 3.8e-39;
Matches 135; Conservative 76; Mismatches 148; Indels 53; Gaps 13;
QY 62 DELNEINQFLGPVEKEFTTEV--DSRKIDQEGKIPDETLEKLSLGLFLGLOVPEEY-- 115
DB 10 DEQKEVRDWL---HGFAADVIRPAAAEWDERETPPVPIQEAQKVGYSLDYFAQQFDF 65
QY 116 -GGIGFSNTMYSRLEGIISMDGSITVTAHQAIGLKGIIAGLAGEEOKAKYLPKL---AS 171
DB 66 PTGLGIPMAN-----EELFWGCDAGIALSIVG--TGLAAGVGLANGTEEQIGTWIPQWGDAN 120
QY 172 GEHIAAFCLTEPASGSDAASIRSRATLSSEKHHYILNGSKVWITNGLANIETVFAKTEV 231


```

Db 121 DVKAAFCSEPOAGSVSMRTRAVYDEAKDEWLNKTKWATNGIANHVHVA---v 177
Qy 232 VDSGSKVDKITAFIVERDFGGYNGKPKDLGRSNTCEVHFENTKIPVENILG---- 287
Db 178 VDSLGSKGH-ASFIVPNTPGUSQOKFKKHGIRASHTAENVLENARIPGSCLLGKKEK 236
Qy 288 -----EVGDGFKV---AMNLSGRFSMGSVVAGLLKRLIENTAEVACTRKOF 332
Db 237 LDERLARERARAGGERVKNAAMATEASRPVAGAMAVGTARAAYEBALEYAKTREQF 296
Qy 333 NKRLSE-----FGLQEKALMAQKAYVMESMTYLTAGMLDQGPDCSIEAAVKVFSS 387
Db 297 GRPIIDNOGVAFOLADWRTSIDAARLLVWR-----ASWYVNGKPFATAEGSMKLPAS 350
Qy 388 EAAQCVSEALQTLGLGVTDRDYPYERILDRTRILLIFEGTNEILRMYYALT 439
Db 351 ETAKKVTAQAVQILGCGGTREYVVERMHRDAIYTFEGTSEIORLVART 402

RESULT 13
US-10-181-319-33
; Sequence 33, Application US/10181319
; Publication No. US20030135032A1
; GENERAL INFORMATION:
; APPLICANT: Lewis, Thomas A.
; APPLICANT: Paszczynski, Andrzej
; APPLICANT: Crawford, Ronald L.
; APPLICANT: Cortese, Marc S.
; APPLICANT: Sebat, Jonathan L.
; TITLE OF INVENTION: Compositions and Methods for Bioremediation
; FILE REFERENCE: IDAHL1940
; CURRENT APPLICATION NUMBER: US/10/181,319
; CURRENT FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: PCT/US01/02386
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/177,251
; PRIOR FILING DATE: 2000-01-20
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Pseudomonas stutzeri
US-10-181-319-33

```

```

Query Match 15.2%; Score 480; DB 12; Length 513;
Best Local Similarity 30.2%; Pred. No. 9.2e-38;
Matches 136; Conservative 94; Mismatches 181; Indels 40; Gaps 12;

Qy 9 RTTAAARACGLVSTANRRLLRTSPVPRAFAKELFLGKIKKK-EVPPFPPEVSQ----- 61
Db 70 RATCARRNRQVFSCHLRLLH-HPKHRRARGVPLGPPRWQLDPSGRLSRLAGTPE 128
Qy 62 --DEL-----NEINQFLGPVEKFTTEVD--SRKIDQEGKIPDETLEKLK----SLGLFG 108
Db 129 ANDMIDFSLPNEQMLVSTVKRFVENELNPLEDEIERTNAIDPSVAEGLKQKAREGLWA 188
Qy 109 LQVPEYGGGLFSNTYSLRGEIISMDGTSVTTLAHOAIG-LKGIILAGTEQKAKYLP 167
Db 189 MHMPQEVGGGLSAVEFCLVNEQIGR---TKDVLARRAFGHVPSILVHCTGSEQREKYLH 244
Qy 168 KLASGEHIAAFLCTEPASGDAASIRSRATLSDDKKHYILNGSKVMTNGGLANITVFA 227
Db 245 AAMRGOIHVSAMSEPEAGSDANGIRT--AVKRDGCEWILNGSKHFSIDADIASYIVTA 302
Qy 228 KTEVDSGSKVDKITAFIVERDFGGYNGKPKDLGRSNTCEVHFENTKIPVENILG 287
Db 303 RSE-----EGISCFLVDRDTPGLELGPQIQQEMMGRGTHQHGFLFFDCRIAPQQLG 353
Qy 288 EVGDGFKVAMNINLSRES-MGSVAVAGLLKRLIENTAEVACTRKQFNKRLSEFGIQRKF 346
Db 354 EPCRGNLSVLGHLNVARLAVGARAVGMAKLLMSVDFAKORSQFGAPIGSFQMVQKML 413

```

```

Qy 347 ALMAQKAYVMESMTYLTAGMLDQGPDCSIEAAVKVFSSAAHQCVSEALQTLGLGY 406
Db 414 ADMQCEIYGARMMLLNAWEIDQG--RDVREKVSMLKLFASEMLGRVADSAVOIFGGMGY 471
Qy 407 TRDYPYERILDRTRILLIFEGTNEILRMYYIA 437
Db 472 CTPLPIERYRDARVRLYDGTSETHRIMIA 502

```

RESULT 14

```

US-10-205-823-4
; Sequence 4, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Womsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-4

```

```

Query Match 15.2%; Score 479; DB 15; Length 415;
Best Local Similarity 30.9%; Pred. No. 8.1e-38;
Matches 137; Conservative 76; Mismatches 195; Indels 36; Gaps 9;

Qy 2 SCCGLFLRTTAAARACGLVSTANRRLLRTSPVPRAFAKELFLGKIKKKVFPPEVSQ 61
Db 4 SCRRFGRALGCLPGRLVLVQTGHRSLTSCIDPSMGLNEE-----QKE---FQKVAF 53
Qy 62 DELNEINFLGPVEKFTTEEV--DSRKIDQEGKIPDETLEKLKSLGLFGLQVPEYGGGLG 119
Db 54 D-----FAAREMAPNNAEWDQKLELFPDVMRKAALQGFEGGVVIQTDVGSG 99
Qy 120 FSNMTYSRLGEIISMD-GSITVTLLAAHQAIKGLKGIILAGTEQKAKYLPKLASGEHTA 178
Db 100 LSRDLTSVIFEALATGCTSTTAYISIHMCAMW-IDSGNEEQHKECPPLCTMEKFASY 158
Qy 179 CLTEPASGDAASIRSRATLSDDKKHYILNGSKVMTNGGLANITFTVFAKTEVVDSDG 238
Db 159 CLTEPGSGDAASLLTSAKKQD--HYILNGSKAFISGAGESDIYVVMCR-----GGPG 211
Qy 239 KDKITAFIVERDFGGYNGKPKDLGRSNTCEVHFENTKIPVENILGEVGDGFKVAMN 298
Db 212 PKGISCIYVEKGTPLSGFKKKEKVGWNSQPTRAVIFEDCAVPVANNRIGSEGQGLIAVR 271

```

306 INQILMAKYVASKAAVDVAGKAVOLLGANGCHADYAVERYRDAKIMELIESTQIHEIQ 365

Db 306 INQILMAKYVASKAAVDVAGKAVOLLGANGCHADYAVERYRDAKIMELIESTQIHEIQ 365
QY 436 IAL 438
Db 366 IAM 368

Search completed: September 3, 2003, 13:39:34
Job time : 64 secs

QY 299 ILNSGRFSMGVAVAGLEKRLIEMTAETACRKNRKLSEGLIQKFAALMAQKAYVMS 358
Db 272 GLNGRINIASCSLGAHAHVILTRDLNVRKQFGEPLASNOYLOFTLADMATRLVAARL 331
QY 359 MYTILTAGMLDQPGFPDGSIEAAMVKVFSSEAAWOCVSEALQILGGLGYTRDYPYERILRD 418
Db 332 MVRNAVAL-OEERKDAVALCSMAKLFATDECAICNOLMGGYGLKDYAVQVVRD 390
QY 419 TRILLIPEGTNEILRMVIALTGLQ 442
Db 391 SRVHILEGSNEVMRILISRLIQ 414

RESULT 15
US-10-238-075-951
; Sequence 951, Application US/10238075
; Publication No. US20030148324A1
; GENERAL INFORMATION:
; APPLICANT: I.N.S.E.R.M.
; TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolated from the cells of Escherichia coli and biological uses of these polynucleotides and of their derivatives
; FILE REFERENCE: BLANDINE
; CURRENT APPLICATION NUMBER: US/10/238,075
; PRIOR FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 0003145
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 1576
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 951
; LENGTH: 376
; TYPE: PRY
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (164)..(164)
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (254)..(254)
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (285)..(285)
; OTHER INFORMATION: unknown
US-10-238-075-951

Query Match 14.4%; Score 452.5; DB 12; Length 376;
Best Local Similarity 31.1%; Pred. No. 2.6e-35;
Matches 113; Conservative 71; Mismatches 152; Indels 27; Gaps 7;
QY 84 SRKIDQGGKIPDETLEKLSGLGLQVPPEYGGGLGFSNTMYSRLGELI-----SMDGSI 138
Db 25 AQNIERQOFIVPDIISCVAGAGYLGASIPQKYGGRGVDSYQLCALHEVMAGVHGSLENLI 84
QY 139 TVTLAAHQATGLKGIIL--AGTEEQAKYLPKLASEHIAAFCITPEASGSDAASIRRA 196
Db 85 TVT-----GMVSTLLQVRGSAQAQAHYLPKLATGELIGALIAITLPEFNGSDLVNVETE- 136
QY 197 TLSDEKKHYILNGSKVITNGGLANITFTVFAKTEVVDSDGSKVDKITAFIVERDFGGVTN 256
Db 137 -LQDGGGWRNLGKKWITLQGTADFXLI-----HCGNQLATVLIDRNTDGTFTI 186
QY 257 GKPEDKLGIRGNTCEVHFENTKIPVENILGEVGDGPKVAMNT-LSNGRFSMGSVVAGLL 315
Db 187 TPLNDMLGRGNLAEHLHFNDCRLKEDALLGPTPGVPLAVNFALNEGRETTCAGSLGIC 246
QY 316 KRLIEMTAETACRKNRKLSEGLIQKFAALMAQKAYVMSYTLTAGMLDQPGFPDC 375
Db 247 QAAVDVAXRYIRQKQKRRLLFSHGIVOHLEFATMLTQTSXQALMCF-SAAEYRETLHPAM 305
QY 376 SIEAAMVKVFSSEAAWOCVSEALQILGGLGYTRDYPYERILRTRILLIPEGTNEILRMV 435

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 13:26:07 ; Search time 30 Seconds
(without alignments)
875.835 Million cell updates/sec

Title: US-09-945-326-2
 Perfect score: 3153
 Sequence: 1 MSGGGLFLRTTAARACRL.....SQILEKRAYICAHPLDRTC 621

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

```

Database :
1: /cgn2_6/pdata1/1aa/5A_COMB.pep.*
2: /cgn2_6/pdata1/1aa/5B_COMB.pep.*
3: /cgn2_6/pdata1/1aa/6A_COMB.pep.*
4: /cgn2_6/pdata1/1aa/6B_COMB.pep.*
5: /cgn2_6/pdata1/1aa/PCFUS_COMB.pep.*
6: /cgn2_6/pdata1/1aa/backfiles1.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	691.5	21.9	390	4	US-09-252-991A-26058	Sequence 26058, A	
2	666	21.1	389	4	US-09-328-352-6380	Sequence 6380, Ap	
3	650	20.6	384	4	US-09-648-004-6	Sequence 6, Appli	
4	623	19.8	387	4	US-09-328-352-6442	Sequence 6442, Ap	
5	618.5	19.6	403	4	US-09-252-991A-27960	Sequence 27960, A	
6	608	19.3	382	4	US-09-328-352-4850	Sequence 4850, Ap	
7	603.5	19.1	419	4	US-09-252-991A-31097	Sequence 31097, A	
8	596	18.9	394	4	US-09-328-352-6801	Sequence 6801, Ap	
9	595.5	18.9	409	4	US-09-364-230-30	Sequence 30, Appl	
10	588.5	18.7	427	4	US-09-364-230-32	Sequence 32, Appl	
11	559.5	17.7	383	4	US-09-328-352-7081	Sequence 7081, Ap	
12	514.5	16.3	402	4	US-09-328-352-4281	Sequence 4281, Ap	
13	501	15.9	424	4	US-09-252-991A-24653	Sequence 24653, A	
14	501	15.9	399	4	US-09-328-352-4360	Sequence 4360, Ap	
15	495	15.7	464	4	US-09-252-991A-33108	Sequence 33108, Ap	
16	472	15.0	282	4	US-09-252-991A-29144	Sequence 29144, A	
17	467.5	14.8	408	4	US-09-328-352-6854	Sequence 6854, Ap	
18	465	14.7	408	4	US-09-328-352-5575	Sequence 5575, Ap	
19	456	14.5	270	4	US-09-364-230-28	Sequence 28, Appl	
20	434.5	13.8	863	4	US-09-252-991A-19574	Sequence 19574, A	
21	422.5	13.4	415	4	US-09-328-352-4699	Sequence 4699, Ap	
22	415.5	13.2	393	4	US-09-252-991A-23105	Sequence 23105, A	
23	394.5	12.5	683	4	US-09-252-991A-26189	Sequence 26189, A	
24	392	12.4	413	4	US-09-252-991A-22542	Sequence 22542, A	
25	390	12.4	998	4	US-09-252-991A-21058	Sequence 21058, A	
26	375.5	11.9	825	4	US-09-328-352-4516	Sequence 4516, Ap	
27	366.5	11.6	399	4	US-09-328-352-5561	Sequence 5561, Ap	

ALIGNMENTS

RESULT 1

```

US-09-252-991A-26058
; Sequence 26058, Application US/09252991A
; Patent No. 6531795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND
; TITLE OF INVENTION: AERUGINOSA FOR DIA
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26058
; LENGTH: 390
; TYPE: prt
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26058

```

Query Match	21.9%;	Score	691.5;	DB	4;	Length	390;
Best Local Similarity	37.8%;	pred. No.	1.2e-61;				
Matches	147;	Conservative	82;	Mismatches	121;	Indels	39;
						Gaps	8;
QY	67	INQFLGPVEKFFTEEV--DSRKIDQEGKIPDETLEKLSGLGLQVPEYGGGLGFQNTM	124				
Db	14	LNLLDSIQFVRESLPHEQVEAETDRIPETAIARMREMLGFLSLPIEAYGGLG-----	68				
QY	125	YSRLGRIISMDGSGITVT-----LAAHQALGLKGIILAGTECEKAYLKPKLAS	171				
Db	69	-----VTMEEEVSIATFELGRTSPAFRSLLTGNTNGIGSGCIVDGTGEOKRRLPLRAS	121				
QY	172	GERHIAAFCLTEPASPDSDAASIRSRATLSEDKHYILNGSKWTTNGGLANIFTVFARTEV	231				
Db	122	GELLSSFCLETPDGSDDAASLTKTAV--RDGEHYVVLNGTKRFTINAPQAGIYVTMART--	177				
QY	232	VSDSGSVKDK--ITAFIVERDEGGVYNGKPEOKLIGRSGNTCEVHPENTKIPVENTILGEV	289				
Db	178	---DPAIRGAGGISAFVBERGTGSLGLHPDRKMGQKGAHTCDVTFDDCRVPASQIUGV	234				
QY	290	-GDGPKVAMNIIINSGRFSMGSVVAGLLKRLIEMTAETAYACTRKNNKRLSEFGLIOEKFAFL	348				
Db	235	EGVGFKTAMKVLIDKGRLHIAACVGVAAERMLEDALRYALERKQFGQPIAEFQLIQAWLAD	294				
QY	349	MAQKAYVMESMTYLTAGMLDOPGFPDPCSTLEAANVKVFSSEAAWQCUSEALQILGLGUYTR	408				
Db	295	SKAEYAAARCWIDAAARQDEG--RDVGTFEASCAKLFASEMCGRVADRAVOIFGGAGYIG	352				

QY 409 DYPVERILRDRTRILLIEFGTNEILRMVIA 437
Db 353 DIGIERFYRDVRLFRLEYEGTTIQIOOLLIA 381

RESULT 2
US-09-328-352-6380
; Sequence 6380, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6380
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6380

Query Match 21.1%; Score 666; DB 4; Length 389;
Best Local Similarity 38.2%; Pred. No. 4.8e-59;
Matches 150; Conservative 79; Mismatches 148; Indels 16; Gaps 8;

QY 58 EVSQDELNETNOFLGVPVEKFFTEEVDSR--KIDQEGKIPDETLEKLKSGLFGLGVPEEY 115
Db 5 QMIRDE-GMLEQLLTIRDVFKNELPRENEVVEKDPPDDIVQQMRRELGFGLTIPEEY 63

QY 116 GGLGFSNTWYSRLG-EIISNDGSITVTAAHQAIGLKIILAGTETEQAKAYLPKLASGEH 174
Db 64 GGLGITMEEEVRVAFELGQTSPARSLIGTNNGIGSSAILDGTTEEOKKYLPRYASGEI 123

QY 175 IAAFCLTEPASGDASAISRATLSBKXHYILNSKVWTNGGLIANFTTFVAKT--EWV 232
Db 124 IGSFCLTEPESGSDAASLKTSAV--KGDFYVLNGTKREITNAPHAATAFTVMARTNPETK 181

QY 233 DSGSVKDKKTAFIVERDFGVNGPKDGKLGIRGNTCEVHFENTKIPVENILGEV-GD 291
Db 182 GSGG-----ISAFLIVEANTEPIITLGIKDQMGGKGSHTCDVIFENRVPSALIGGVEGV 236

QY 292 GFKYAMNILNSGRFSMGVSVAGLLKRLIEMTAICTRKOFNRKLRSEGLIOEKFAIMQAQ 351
Db 237 GFKTAMKVLVDKGRHLHCAYSVGVAERMLNDALNYAIERKOFQGQPFIANFOLIQA MLADSKA 296

QY 352 KAYWESMTYLTAGMLDPQFPDCSTEAAMVKVESFEAAWCVCSEAQLQILGGLGYTRDYP 411
Db 297 EIIAAKCMVLDAARRONG--ENISTEPACAKMFATEMCGRVADCRCVQHGGAGYISEYA 354

QY 412 YERILRDRTRILLIEFGTNEILRMVIALTLGHQA 444
Db 355 IERYRDVRLFRLEYEGTTQVOQLLIANKNMIREA 387

RESULT 3
US-09-648-004-6
; Sequence 6, Application US/09648004
; Patent No. 6498242
; GENERAL INFORMATION:
; APPLICANT: CHEN, QIONG
; APPLICANT: THOMAS, STUART
; APPLICANT: NAGARAJAN, VASANATHA
; TITLE OF INVENTION: BIOLOGICAL METHOD FOR THE PRODUCTION OF ADIPIC ACID AND
; TITLE OF INVENTION: INTERMEDIATES
; FILE REFERENCE: CL-1341-A
; CURRENT APPLICATION NUMBER: US/09/648.004
; CURRENT FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/252,553
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Microsoft Office 97

Qy 225 VFATKTEVSDGSKVKITAFIVERDGGVTNGKPKDKLGIRGNTCEVHFENKIPVEN 284
Db 171 VMARTN-FDIKGA--SGISAFIVDSOTPGISLGRKDKMGQKGVHFTCDVIFENCRI PASA 227
Qy 285 ILGEV-COGFKVAMNINSGRFSMGSVVAGLLKRLIEMTAEXACTRKQFNKRLSEFGLIQ 343
Db 228 LIGGVEGVGFKTAMKLDKGRHLIAALSVAATRDLDLSLVAIERKQFGQPIAEFOLIQ 287
Qy 344 EKFLMAQKAYVNESMTYLTAGMLDQPGFDCSIEAAVKVFSSEAAWOCVSEALQILGG 403
Db 288 AMLADSKAEIYAAKCHV-LDAARLRDAG-QNVSTEASCAKMEATEMCGRVADRCVQIHGG 345
Qy 404 LGWTRDYPYERILDRILLIFEGTNEILRWYIA 437
Db 346 AGVISEYAIERFYRDLRLVEGTQIQOVIIA 379

RESULT 5
US-09-252-991A-27960
; Sequence 27960, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: 107196.136
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27960
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27960

Query Match 19.6%; Score 618.5; DB 4; Length 403;
Best Local Similarity 37.5%; Pred. No. 3.5e-54;
Matches 149; Conservative 71; Mismatches 148; Indels 29; Gaps 8;

Qy 52 EVFPPEVSDQELNEINQFLCPVEKFEETEEVDNR--KIDQEGKIPDETLEKLSGLFLG 109
Db 15 ESMYPSLNFALGETIDMLRDQVRGFAAELOPRAAQIDQDNQFPMDMWRKFGEMGLGI 74
Qy 110 QVPEYGGGLFSNTWYSLRGEIIS-MDGSITVTTLAAHQAIKGLIAGTEEQKAKYLPK 168
Db 75 TVDEYGGSGALGYLAHVAVMEIISRASASVALSYGASHNLVCVQIKRNGNAEQARYLPA 134
Qy 169 LASGEHIAAFLCTEPASGSDAASRSTRATLSSEDKKHYILNGSKVWITNGGLANIFTVFAK 228
Db 135 LVSGEHIGALAMSEPNAGSDVSMKLRADRVGDR--FVLNGSKWITNGPDATYVYIAK 192
Qy 229 TEVVDSDGSKVKD-----ITAFIVERDGGVTNGKPKDKLGIRGNTCEVHFENKIPVEN 284
Db 193 TDA-----DRGAHGITAFIVERDQWGSRGPKLQKLGMRGNTCELIQFQVPEEN 244
Qy 285 ILGEVGRKRVAMNINSGRFSMGSVVAGLLKRLIEMTAEXACTRKQFNKRLSEFGLIOE 344
Db 245 VLCAVNGGVKVLMSGLDYERVLVLSGGFVGIMQACMDVVVPIYIHDRRQFGQSIGEFOLVG 304
Qy 345 KF-----ALMAQKAYVNESMTYLTAGMLDQPGFDCSIEAAVKVFSSEAAWOCVSEALQ 399
Db 305 KVADMVYALNASRAYL-----YAAACDRG--EITRKAAGVILYSERATQMALDAIQ 357
Qy 400 ILGGLGYTRDYPYERILDRILLIFEGTNEILRWYI 436
Db 358 ILGGNGYINEFFPTGRLLRDAKLYEAGTSEIRRMILI 394

RESULT 6

US-09-328-352-4850
; Sequence 4850, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: GTC99-03PA
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4850
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4850

Query Match 19.3%; Score 608; DB 4; Length 382;
Best Local Similarity 37.5%; Pred. No. 3.8e-53;
Matches 142; Conservative 71; Mismatches 154; Indels 12; Gaps 6;

Qy 63 ELNEINQFLCPVEKFEETEE---VDSRKIDQEGKIPDETLEKLSGLFLGLOVPEEYGGGLG 119
Db 9 QFTEEQLLIRDMAKSFAQEOIKPNASDNRDGTTPKRETLTQMQLGFMGMLVSEEWGSD 68
Qy 120 FSNWY--SRIGEISMDGSIYVTLAAHQAIKGLIAGTEEQKAKYLPKLAGSHEHIAAF 178
Db 69 TGNLAYVALLEEVAADGATSTIMSVHNSVGVCPILKFGTDEQKRELAQGENIGAF 128
Qy 179 CLTEPASGSDAASRSTRATLSSEDKKHYILNGSKVWITNGGLANIFTVFAKTEVVDSDGV 238
Db 129 ALTEPHGSDAAAKTRAV--KGDYDYLNGAKQFTSGNNAGVIIVFAVTD---PSAG 182
Qy 239 KDKITAFIVERDGGVTNGKPKDKLGIRGNTCEVHFENKIPVENILGEVGDGPKVANN 298
Db 183 KKGISAFVLPRETPGVEVIRVEKLGHASDTCOIALTDVRIHKSMLGKEGEGKLIALA 242
Qy 299 ILNSGRFSMGSVVAGLLKRLIEMTAEXACTRKQFNKRLSEFGLIOEKFALMAQKAYVNES 358
Db 243 NLEGGRIGIAQAVGLAARALAEATRYAKERITFGKPIFEHOTIAFRLASMAATEIAARQ 302
Qy 359 MYLTAGMLDQPGFDCSIEAAVKVFSSEAAWOCVSEALQILGGLGYTRDYPYERILRD 418
Db 303 LVHY-AARLKEAGQP--CLNEASMAKLFASEMTERVCSSALQVFGGYGLRDEPIRIYRD 360
Qy 419 TRILLIFEGTNEILRWYIA 437
Db 361 ARICQIYEGTSDIQRVLIA 379

RESULT 7
US-09-252-991A-31097
; Sequence 31097, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: 107196.136
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31097
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31097

Query Match 19.1%; Score 603.5; DB 4; Length 419;

Best Local Similarity 38.7%; Pred. No. 1.3e-52; Mismatches 139; Indels 9; Gaps 6; Matches 136; Conservative 67;

Qy 88 DOBKIPDETLKLSGLGLOVPEYGGIIFSNMYS-RLGEIISMDGSTVITLAAHQ 146
 Db 74 DREHFPFAEIREMADLGLFGLMVLPEWGAQOCHLAYAMALBEIAAGDGACSTIMSVHN 133
 Qy 147 AIGLIGIILAGTEOKAKYLPKLASGHIAAFCLTEPASGSDAASIRSRATLSSEKHHYI 206
 Db 134 SVGCMPIHFGSABOKERERPLAQGSMGLGAFALTEPOAGSDASFLLKTRA--RRDGDHYV 191
 Qy 207 LNSKVVITNGGLANITFVFAKTEVVDSDGSKDKITAFIVERDFGVTKGPKEDKLGIR 266
 Db 192 LNAKQFITSAGSHAGMIVFA---VTDPDAG-KRGISAFIVPTDGYEVVRIEDKLGQH 247
 Qy 267 GSNCEVHFENTKIPVENILGEVGDGFKVAMNINLSGRFSMGSVVAGLLKRLIEMTAEYA 326
 Db 248 ASDTCOLAFNDLRIPATLRLGEEGEGYRIALANLEGGRIATAAGVGMARAAFEAARDYA 307
 Qy 327 CTRKQFNKRLSEFGLIOEKFAKMAKAYVMESMTYLTAGMDOPGPDGCSIEAAMVAVFS 386
 Db 308 HERETFGPIIHOAVAFRLADMATRIAVARQMVH-HAASLRAGLP-CLTEASMAKLFA 365
 Qy 387 SEAAWQCSEALQILGLGIVTRDYPYERILRDRTRILLIFETNEILRMVIA 437
 Db 366 SEMAEVCSAALQILGGYGLKDFPVERIYRDRVRCQIYEGTSDVRLVIA 416

RESULT 8
 US-09-328-352-6801
 ; Sequence 6801, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 6801
 ; LENGTH: 394
 ; TYPE: PRT
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-6801

Query Match 18.9%; Score 596; DB 4; Length 394;
 Best Local Similarity 37.5%; Pred. No. 6.6e-52;
 Matches 140; Conservative 73; Mismatches 138; Indels 22; Gaps 9;

Qy 74 VERKFFTEVD--SRKIDQEGKIPDETLEKLSGLGLOVPEYGGIIFSNMYSRLGEI 131
 Db 25 VAAFAKAIPIAQVDDQDKFPAHLWKFGDMGLLGMTVSEYGG---ANNGY--LAHI 79
 Qy 132 ISM-----DGSITVTLAAHQALGKILAGTEOKAKYLPKLASGEHIAAFCLTEPAS 185
 Db 80 IAMQETISRAAATGLSYGAHSLNVCNINRNGEQOKYLPKLISGEVVGALAMSEPN 139
 Qy 186 GSDAASIRSRATLSSEKHHYILNSKVVITNGGLANITFVFAKTEVVDSDGSKDKITAF 245
 Db 140 GSDVSMKLRAGOKGD--HVLNLSKVVITNGGDADVLVYAKT---DPQAGPKG-WTAF 193
 Qy 246 IVERDFGVVTKGPKEDKLGIRGNTCEVHFENTKIPVENILGTEVGGFKVAMNINLSGRF 305
 Db 194 LIEKMGKFGSHNLDKLGMRGNTYPLFFDNVVEPAENLVGGVGVKVLMSGLDYERA 253
 Qy 306 SMGSVAGLLKRLIEMTAETVACTRKQFNKRLSEFGLIOEKFAKMAKAYVMESMTYLTAG 365
 Db 254 VLSAGPLGIMADCLDVVPIYHOREFGQALGFEQLMOGLADMYSTWLACALVAVGA 313
 Qy 366 MLDOPGPDGSI--EAAWQVVFSEAAWQCSEALQILGLGYTRDYPYERILRDRTRILL 423
 Db 314 ACDAKH-DRSLRKDAASAILYAAEKATWAMAGEAIOITLGGNGYINFEPPAGRLWRDAKLYE 372

Qy 424 IFEGTNEILRMVYI 436
 Db 373 IGATSEIRRMILI 385
 RESULT 9
 US-09-364-230-30
 ; Sequence 30, Application US/09364230
 ; Patent No. 6348339
 ; GENERAL INFORMATION:
 ; APPLICANT: Cahoon, Rebecca E.
 ; APPLICANT: Hitz, William D.
 ; APPLICANT: Kinney, Anthony J.
 ; APPLICANT: Rafalski, J. Antoni
 ; TITLE OF INVENTION: Enzymes Involved in Degradation of Branched-Chain Amino Acids
 ; FILE REFERENCE: BB-1178
 ; CURRENT APPLICATION NUMBER: US/09/364,230
 ; CURRENT FILING DATE: 1999-07-29
 ; EARLIER APPLICATION NUMBER: 60/094,990
 ; EARLIER FILING DATE: July 31, 1998
 ; NUMBER OF SEQ ID NOS: 34
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 30
 ; LENGTH: 409
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 US-09-364-230-30

Query Match 18.9%; Score 595.5; DB 4; Length 409;
 Best Local Similarity 38.5%; Pred. No. 8e-52;
 Matches 150; Conservative 64; Mismatches 143; Indels 33; Gaps 13;

Qy 62 DELNEINQFLGPVEKPFTEEV--DSRKIDQEGKIP-DETELEK-L-SIGLFGLOVPEYGG 117
 Db 32 DDTQE--QPKESVHKFAQETIAPHAADAIDASNHPKDVNLWKLMDGDFNLHGLTAPEYGG 89
 Qy 118 LGSNTMYS--RLGEIISMDGSTVITLAAHQALGKILAGTEOKAKYLPKLASGEHI 175
 Db 90 MGL-GYMIHCIAEMEEINRSGSVGLSYSAHSLNCLINOLVRHSGSPAOKLKLPLKLTGEHV 148
 Qy 176 AAFCLTEPASGSDAASIRSRATLSSEKHH--YILNLSKVVITNGGLANITFVFAKTEV-V 232
 Db 149 GALAMSEPNSGSDVSMCKRA---ERKVGYYINGNKMCTNGSPSAQTUGVYAKDITA 204
 Qy 233 DSDGSVDKKTAFIVERDFGVTKGPKEDKLGIRGNTCEVHFENTKIPVENILGEYGDG 292
 Db 205 GSKG-----ITAFIIEKMGAGFSTAOKLDKLMRGSDTCELVEFCFPHENVLGERGKG 259
 Qy 293 FKVAMNINLSGRFSMGSVVAGLLKRLIEMTAETVACTRKQFNKRLSEFGLIOEKFAKMAOK 352
 Db 260 VYVMMSGLNLERFVLAAGPSALMQACLDVAVLYVRQREQFGRPIGEFQFIRGLADMYTS 319
 Qy 353 AYVMESMTYLTA---GMLDOPGPDGCSIEAAMVVFSSAAWQCSEALQILGLGYT 407
 Db 320 LQSSRSFVYSVARCDNKGVDK---KDC---AGVILFAERATQVALQIOLGNGYI 372
 Qy 408 RDYPERILRDRTRILLIFETNEILRMVIA 437
 Db 373 NEYPTARLLRDAKLFEGIPGTSEIRRMIIA 402

RESULT 10
 US-09-364-230-32
 ; Sequence 32, Application US/09364230
 ; Patent No. 6348339
 ; GENERAL INFORMATION:
 ; APPLICANT: Cahoon, Rebecca E.
 ; APPLICANT: Hitz, William D.
 ; APPLICANT: Kinney, Anthony J.
 ; APPLICANT: Rafalski, J. Antoni
 ; TITLE OF INVENTION: Enzymes Involved in Degradation of Branched-Chain Amino Acids
 ; FILE REFERENCE: BB-1178

[illegible]

03-09-232-991A-24653
; Sequence 24653, Application US/09252991A

Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 24653
LENGTH: 424
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24653
Query Match 15.9%; Score 501; DB 4; Length 424;
Best Local Similarity 33.8%; Pred. No. 3.6e-42;
Matches 127; Conservative 69; Mismatches 158; Indels 22; Gaps 8;
QY 70 FLGPVEKFFTEV--DSRKIDQEGKIPDETLEKLSGLFLGLQVPEYGGGLGFNTMYSR 127
DB 60 FRDSVRKFELEQEA VPFHWEKDGHVDRALWNKAGEAGMCLCSHLPEYGGMG-ADFLYSA 118
QY 128 LG-ETISMDGSTVTLAAHQAGLKGIIILAGTEEOKAKYLPKLAGSEHIAAFCLTEPASG 186
DB 119 VVIEETGRAGLTGIGFSLHSDIAAPYILHYSEALKRKLPLKLVSGEMVAALMTEPGAG 178
QY 187 SDAASIRATLSSEDKHYILNGSKVITNGGLANITVFATKTEVVDSDGSKDKITAFI 246
DB 179 SLOGVKITAVLDGB--YVINGSKTFTNGFLADLVIVAKT---DPKAGAKG-ISLFV 232
QY 247 VERDFGGVNGPKDLGIRGNTCEVHFENTKIPVENILGVDGFKVAMNLSGRFS 306
DB 233 VEAGTGFSGKRLEKVGKMAQDTSELFDDVRIPRENLLGKDGOGFIYLMQELPQERLT 292
QY 307 MGSVAGLLKRLIEMTAETACTRKOFNKRSLSEFLIQKLFALMAQ-----RAYVMESMTY 361
DB 293 VGIGALASAEALOWTLDTYREKAFGKSVADFQNTFRKLAEMATEIQVGRVFDRCMEQ 352
QY 362 LTAGMLDQPGFPCDSTEAMVVFSEAAWQCVSEALQILGLGYTRDYPYERILDRFI 421
DB 353 HLQCKLDVP-----TAAMCKYWTITDLOCKVLDCEVQJHGGYGFEMWEPYVARAWADRV 405
QY 422 LLIFEGTNEILRMVIA 437
DB 406 QRIYAGTNEIMKEIIA 421
RESULT 14
US-09-328-352-4360
Sequence 4360, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 4360
LENGTH: 399
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-4360
Query Match 15.7%; Score 495; DB 4; Length 399;
Best Local Similarity 32.3%; Pred. No. 1.3e-41;
Matches 124; Conservative 70; Mismatches 162; Indels 28; Gaps 10;

QY 66 EINQFLGPVPEKFFTEVDSR--KIDQEGKIPDETLEKLSGLFLGLQVPEYGGGLGFNT 123
DB 25 DLELFRDNFKRPMNEHIAPHYDQWREGIMPRVSWSQLGENGFLCVDVPEYGGVGP-T 83
QY 124 MYS-RLGEIISMDG--SITVTTLAAHQAGLKGIIILAGTEEOKAKYLPKLAGSEHIAAFCL 180
DB 84 YYSMLVSEASARAGFCALSTASISEIAAPYILHIGTEEOKQYWLPMKMTGEVVGALGM 143
QY 181 TEPASGSDAASIRATLSSEDKHYILNGSKVITNGGLANITVFATKTEVVDSDGSKVD 240
DB 144 TEPGAGSDLOSMTSAILQDD--HYLLNGSKTFTNGFLADLVIVAKT-----DPOARA 196
QY 241 K-ITAFIVERDFGGVNGPKDLGIRGNTCEVHFENTKIPVENILGVDGFKVAMNI 299
DB 197 KGVSIILLVDTHLEGFPKGTNLDKIGLSQDTSELFDDNVKVPKNQLLQGLGAGGQFAYLMQE 256
QY 300 LNSGRFSMGSVAGLLKRLIEMTAETACTRKOFNKRSLSEF-----LIQEKFALMAOKAY 354
DB 257 LPRETAIASTAVGAIRGAIDLATAVYKERRHAFGQPIISOQFQNTFRVLAQAKIDELATTAF 316
QY 355 VMESMTYLTAGMLDQPGFPCDSTE-AMVVFSEAAWQCVSEALQILGLGYTRDYPYE 413
DB 317 YERNVALYQBGKLD-----VETAAALKSFSTDMQKVNADNLLQLFGGIGYMTYEPIS 368
QY 414 RIIRDTRILLIPFGTNEILRMVIA 437
DB 369 RFFVDARIQRIYGGTNEIMKEIVA 392
RESULT 15
US-09-252-991A-33108
Sequence 33108, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 33108
LENGTH: 464
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-33108
Query Match 15.7%; Score 495; DB 4; Length 464;
Best Local Similarity 32.9%; Pred. No. 1.7e-41;
Matches 144; Conservative 77; Mismatches 179; Indels 38; Gaps 13;
QY 12 AAARACRGLVSTANRRLLRTSPVY-RAFAKELFLGKIKKKEVEFPF---EVSODELNE- 66
DB 44 AALRAQRCGRGAA--RRGRETFPSIQRALGAE-----OVTPMDRFLTQEQDMLVEA 93
QY 67 INQFLGPVPEKFFTEVDSRKKIDQEGKIPDETLEKLSGLFLGLQVPEYGGGLG--- 119
DB 94 VRSF---VEKELLPHEDD--VDRADAVSPELAAQIRGALAAAGFAFNMPPEVGGGLDY 148
QY 120 FNTMYSLGETIISMDGSTVTLAAHQAGLKGIIILAGTEEOKAKYLPKLAGSEHIAAF 179
DB 149 LSQLVER--ELSKVSWALHVFVAPRSK-----ILMACTGOLGDLPLPCVOGETDCFA 201
QY 180 LPEPASGSDAASIRATLSSEDKHYILNGSKVITNGGLANITVFATKTEVVDSDGSKV 239
DB 202 LPEPASGSDANSIKTRAV--RGDFAVINGSKHFTSHAGHADFAIVFAVTDYSYEHNGRR 259
QY 240 DKITAFIVERDFGGVNGPKDLGIRGNTCEVHFENTKIPVENILGVDGFKVAMNI 299

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
OM protein - protein search, using sw model
Run on: September 3, 2003, 13:24:27 ; Search time 108 seconds
(without alignments)
1483.802 Million cell updates/sec

Title: US-09-945-326-2
Perfect score: 3153
Sequence: 1 MSCGGLFLRTTAARACRGL.....SQILEKRAYICAHPLDRTC 621

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phase:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rviro:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3153	100.0	621	4	Q9H845
2	3149	99.9	621	4	Q8WXX3
3	2704	85.8	625	11	Q8JZN5
4	2703	85.7	625	11	Q8BK76
5	2700	85.6	625	11	Q8C0B5
6	2526	80.1	498	4	Q9H9W4
7	1318	41.8	655	6	Q8HXV7
8	1308	41.5	633	4	Q8WU10
9	1295.5	41.1	655	5	Q9V8T1
10	1210.5	38.4	613	5	Q19057
11	954	30.3	188	4	Q9BUX5
12	821	26.0	594	16	Q9K784
13	815.5	25.9	594	16	Q32176
14	789.5	25.0	594	16	Q8ENT4
15	734	23.3	169	11	Q9CZC5
16	719.5	22.8	382	16	Q8EM49

17	709.5	22.5	381	16	Q8R674
18	699	22.2	382	16	Q9AB07
19	699	22.2	565	16	Q8ES29
20	695.5	22.1	380	16	Q8RC91
21	691.5	21.9	384	16	Q9I391
22	690	21.9	379	16	Q9K6D0
23	688	21.8	380	16	Q8FUX4
24	682	21.6	381	16	Q8YDG3
25	681	21.6	380	2	Q8VOK1
26	678	21.5	381	16	Q9A7C6
27	672.5	21.3	375	2	Q9AHX9
28	667	21.2	379	16	Q8EM50
29	659	20.9	384	16	Q8FX67
30	658	20.9	380	16	Q9K6D1
31	655.5	20.8	389	2	Q8KZQ2
32	654.5	20.8	379	2	Q8RM16
33	652	20.7	387	16	Q9RUX5
34	651.5	20.7	384	2	Q33772
35	650.5	20.6	636	16	Q8R5V8
36	650	20.6	384	2	Q9F7E6
37	650	20.6	384	16	Q8ESF1
38	648	20.6	379	2	P97088
39	642.5	20.4	385	16	Q8XR24
40	642	20.4	658	16	Q9KYS2
41	633.5	20.1	379	16	Q8X124
42	629.5	20.0	381	16	Q9RVV0
43	628.5	19.9	385	16	Q9RIQ5
44	628	19.9	456	16	Q8YB77
45	626	19.9	380	16	Q8UAA0

ALIGNMENTS

RESULT 1

ID	Q9H845	PRELIMINARY;	PRT;	621 AA.
AC	Q9H845;			
DT	01-MAR-2001 (TREMBlrel. 16, Created)			
DT	01-MAR-2001 (TREMBlrel. 16, Last sequence update)			
DT	01-MAR-2003 (TREMBlrel. 23, Last annotation update)			
DE	Hypothetical protein FLJ13950.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,			
RA	Wagatsuma M., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,			
RA	Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,			
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saio K.,			
RA	Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,			
RA	Ninomiya K., Iwayanagi T.;			
RT	NEDO human cDNA sequencing project.;			
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RC	SEQUENCE FROM N.A.			
RC	TISSUE=Lung, and Uterus;			
RA	Strausberg R.;			
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AK024012; BAB14775.1;			
DR	EMBL; BC013354; AAH13354.1;			
DR	EMBL; BC007970; AAH07970.1;			
DR	HSSP; Q06319; 1BUC.			
DR	InterPro; IPR006089; Acyl-CoA-dh.			
DR	InterPro; IPR006090; Acyl-CoA-dh_C.			
DR	InterPro; IPR006091; Acyl-CoA-dh_M.			
DR	InterPro; IPR006092; Acyl-CoA-dh_N.			
DR	InterPro; IPR005829; Sug_transporter.			
DR	Pfam; PF00441; Acyl-CoA-dh; 1.			
DR	Pfam; PF02770; Acyl-CoA-dh_M; 1.			

Q8r674 fusobacteri
Q9ab07 caulobacter
Q8es29 oceanobacil
Q8rc91 thermoaer
Q9i391 pseudomonas
Q9k6d0 bacillus ha
Q8fux4 brucella su
Q8ydg3 brucella me
Q8vqk1 brucella ab
Q9a7c6 caulobacter
Q9ahx9 pseudomonas
Q8em50 oceanobacil
Q8fx67 brucella su
Q9k6d1 bacillus ha
Q8kzq2 butyrivibri
Q8rm16 clostridium
Q9rux5 deinococcus
Q937t2 acinetobact
Q9f7e6 acinetobact
Q8esf1 oceanobacil
P97088 clostridium
Q8xr24 ralstonia s
Q9kys2 streptomyce
Q8x124 clostridium
Q9rvv0 deinococcus
Q9riq5 streptomyce
Q8yb77 brucella me
Q8uaa0 agrobacteri

```
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
KW Hypothetical protein.
SQ SEQUENCE 621 AA; 68760 MW; 064BCE0378877F54 CRC64;

Query Match 100.0%; Score 3153; DB 4; Length 621;
Best Local Similarity 100.0%; Pred. No. 5.1e-214; Indels 0; Gaps 0;
Matches 621; Conservative 0; Mismatches 0;

QY 1 MSGCGLFLRTAAARACGLVSTANRRLLRTSPVRAFAKELFLGKIKKKEVFPPEVS 60
Db 1 MSGCGLFLRTAAARACGLVSTANRRLLRTSPVRAFAKELFLGKIKKKEVFPPEVS 60
QY 61 QDELNEINQFLGPVEKFFTEVDNRKIDQEGKIPDETLEKLSGLGFLQVPEEYGLGF 120
Db 61 QDELNEINQFLGPVEKFFTEVDNRKIDQEGKIPDETLEKLSGLGFLQVPEEYGLGF 120
QY 121 SNTMYSRLGELIISMDGSIITVLAHQAIIGLKGIILAGTEEQAKYLPKLASGEHIAAFCL 180
Db 121 SNTMYSRLGELIISMDGSIITVLAHQAIIGLKGIILAGTEEQAKYLPKLASGEHIAAFCL 180
QY 181 TEPASGSDAASIRSRATLSDEKHHYILNGSKVITNGGLANITVFAKTEVVDSDGSKVD 240
Db 181 TEPASGSDAASIRSRATLSDEKHHYILNGSKVITNGGLANITVFAKTEVVDSDGSKVD 240
QY 241 KITAFIVERDPGGVNGKPEDKLGIRGNTCEVHFENTKIPVENILGEVGDGFKVAMNII 300
Db 241 KITAFIVERDPGGVNGKPEDKLGIRGNTCEVHFENTKIPVENILGEVGDGFKVAMNII 300
QY 301 NSGRFSMGVSVAGLLKRLIEMTAETACTRKQFNKRLSEFGLIOEKFALMAQKAYVMSMT 360
Db 301 NSGRFSMGVSVAGLLKRLIEMTAETACTRKQFNKRLSEFGLIOEKFALMAQKAYVMSMT 360
QY 361 YLTAGMLDOPGPPDCSIEAAWKVFSSEAAWQCVSEALQILGGIGYTRDYPYERILRDTR 420
Db 361 YLTAGMLDOPGPPDCSIEAAWKVFSSEAAWQCVSEALQILGGIGYTRDYPYERILRDTR 420
QY 421 ILLIFEGTNEILRMVYALTGLQHAGRIITRIHELKQAKYSTVMDTVGRRLRDSLGRTVD 480
Db 421 ILLIFEGTNEILRMVYALTGLQHAGRIITRIHELKQAKYSTVMDTVGRRLRDSLGRTVD 480
QY 481 LGLTGNHGVVHPSLADSANKFEENTYCFGRVTETLLRFGKTIIMEEQVLKRVANILINL 540
Db 481 LGLTGNHGVVHPSLADSANKFEENTYCFGRVTETLLRFGKTIIMEEQVLKRVANILINL 540
QY 541 YGMTAVLSRASRSIRIGLRNHDHEVLANTFCVEAYLQNLFSLSQLDKYAPENLDEQIKK 600
Db 541 YGMTAVLSRASRSIRIGLRNHDHEVLANTFCVEAYLQNLFSLSQLDKYAPENLDEQIKK 600

RESULT 2
Q8WXX3 PRELIMINARY; PRT; 621 AA.
AC Q8WXX3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Very-long-chain acyl-CoA dehydrogenase VLCAD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang W., Zhang J., Li N., Wan T., Chen T., Zhang M., Cao X.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF327351; AAL56011.1; -.
```

```
DR InterPro; IPR006089; Acyl-CoA_dh.
DR InterPro; IPR006090; Acyl-CoA_dh_C.
DR InterPro; IPR006091; Acyl-CoA_dh_M.
DR InterPro; IPR006092; Acyl-CoA_dh_N.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00441; Acyl-CoA_dh_1.
DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
SQ SEQUENCE 621 AA; 68768 MW; 929D7C97A37AC08F CRC64;

Query Match 99.9%; Score 3149; DB 4; Length 621;
Best Local Similarity 99.8%; Pred. No. 9.7e-214; Indels 0; Gaps 0;
Matches 620; Conservative 0; Mismatches 1;

QY 1 MSGCGLFLRTTAAARACGLVSTANRRLLRTSPVRAFAKELFLGKIKKKEVFPPEVS 60
Db 1 MSGCGLFLRTTAAARACGLVSTANRRLLRTSPVRAFAKELFLGKIKKKEVFPPEVS 60
QY 61 QDELNEINQFLGPVEKFFTEVDNRKIDQEGKIPDETLEKLSGLGFLQVPEEYGLGF 120
Db 61 QDELNEINQFLGPVEKFFTEVDNRKIDQEGKIPDETLEKLSGLGFLQVPEEYGLGF 120
QY 121 SNTMYSRLGELIISMDGSIITVLAHQAIIGLKGIILAGTEEQAKYLPKLASGEHIAAFCL 180
Db 121 SNTMYSRLGELIISMDGSIITVLAHQAIIGLKGIILAGTEEQAKYLPKLASGEHIAAFCL 180
QY 181 TEPASGSDAASIRSRATLSDEKHHYILNGSKVITNGGLANITVFAKTEVVDSDGSKVD 240
Db 181 TEPASGSDAASIRSRATLSDEKHHYILNGSKVITNGGLANITVFAKTEVVDSDGSKVD 240
QY 241 KITAFIVERDPGGVNGKPEDKLGIRGNTCEVHFENTKIPVENILGEVGDGFKVAMNII 300
Db 241 KITAFIVERDPGGVNGKPEDKLGIRGNTCEVHFENTKIPVENILGEVGDGFKVAMNII 300
QY 301 NSGRFSMGVSVAGLLKRLIEMTAETACTRKQFNKRLSEFGLIOEKFALMAQKAYVMSMT 360
Db 301 NSGRFSMGVSVAGLLKRLIEMTAETACTRKQFNKRLSEFGLIOEKFALMAQKAYVMSMT 360
QY 361 YLTAGMLDOPGPPDCSIEAAWKVFSSEAAWQCVSEALQILGGIGYTRDYPYERILRDTR 420
Db 361 YLTAGMLDOPGPPDCSIEAAWKVFSSEAAWQCVSEALQILGGIGYTRDYPYERILRDTR 420
QY 421 ILLIFEGTNEILRMVYALTGLQHAGRIITRIHELKQAKYSTVMDTVGRRLRDSLGRTVD 480
Db 421 ILLIFEGTNEILRMVYALTGLQHAGRIITRIHELKQAKYSTVMDTVGRRLRDSLGRTVD 480
QY 481 LGLTGNHGVVHPSLADSANKFEENTYCFGRVTETLLRFGKTIIMEEQVLKRVANILINL 540
Db 481 LGLTGNHGVVHPSLADSANKFEENTYCFGRVTETLLRFGKTIIMEEQVLKRVANILINL 540
QY 541 YGMTAVLSRASRSIRIGLRNHDHEVLANTFCVEAYLQNLFSLSQLDKYAPENLDEQIKK 600
Db 541 YGMTAVLSRASRSIRIGLRNHDHEVLANTFCVEAYLQNLFSLSQLDKYAPENLDEQIKK 600
QY 601 VSOQILEKRAYICAHPLDRTC 621
Db 601 VSOQILEKRAYICAHPLDRTC 621

RESULT 3
Q8JZN5 PRELIMINARY; PRT; 625 AA.
AC Q8JZN5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein (Similar to acyl-CoA dehydrogenase).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC031137; AAH31137.1; -
DR EMBL; BC032213; AAH32213.1; -
DR EMBL; BC033277; AAH33277.1; -
DR InterPro; IPR006089; Acyl-CoA.dh.
DR InterPro; IPR006090; Acyl-CoA.dh.C.
DR InterPro; IPR006091; Acyl-CoA.dh.M.
DR InterPro; IPR006092; Acyl-CoA.dh.N.
DR Pfam; PF00441; Acyl-CoA.dh; 1.
DR Pfam; PF02770; Acyl-CoA.dh.M; 1.
DR Pfam; PF02771; Acyl-CoA.dh.N; 1.
DR PROSITE; PS00072; ACYL_COA_DH.1; 1.
DR PROSITE; PS00073; ACYL_COA_DH.2; 1.
KW Hypothetical protein.
SQ SEQUENCE 625 AA; 68707 MW; 4F06FFBFD82F022 CRC64;

Query Match 85.8%; Score 2704; DB 11; Length 625;
Best Local Similarity 85.7%; Pred. No. 2.6e-182;
Matches 534; Conservative 38; Mismatches 47; Indels 4; Gaps 1;

QY 1 MSGCGLFLR-----TTAARACRGLVSTANRRLLRTSPVRAFAKELFLGKIKKKEVFPF 56
DB 1 MSGCVLLSRGATAAAAAAASRVLRFTARRRPLHTSLQSCFAKELFLGNKQGVFPF 60
QY 57 PEVSQDELNETNQLPGVPEKFFTEVDSDRKIDQEGKIPDETLEKLSLGLFGLQVPEYG 116
DB 61 PEVSQHELSEINQFVPLEKFFTEVDSDRKIDQEGKIPDETLEKLSLGLFGLQVPEYG 120
QY 117 GLGFSNTMYSLRGEIISMDGSIITVTLAAHQAIKGLIILAGTBEQAKYLPKLASGEHIA 176
DB 121 GLGLSNTMYARLGEIISLDSITVTLAAHQAIKGLIILVGNDEQAKYLPKLSGEHIA 180
QY 177 AFLCTEPASGSDAASRATLSLSEDKKHYILNGSKVWITNGGLANFTVFAKTEVVDSDG 236
DB 181 AFLCTEPASGSDAASITQTRATLSLSEDKKYFILNGSKVWITNGGLANFTVFAKTEVVDSDG 240
QY 237 SVKDKITAFIVERDFGGVTKGPKEDKLGIRGNTCEVHFENTKIPVENILGEGVGDFKVA 296
DB 241 SKTDKMTAFIVERDFGGITNGKPKEDKLGIRGNTCEVHFENTRVPVENVLGEGVGDFKVA 300
QY 297 MNLNSGRFSMGSVVAGLLKRLIEMTAETAYACTRKQFNKRLSEFGLIQEKFALMAQKAYVM 356
DB 301 MNLNSGRFSMGSAVAGMLKKLIELTAETAYACTRKQFNRLSEFGLIQEKFALMAQKAYVM 360
QY 357 ESMYTLTAGMLDQPGFPDCSIEAAMVKVFSSEAAQCVCSEALQILGGLGYTRDYPYERIL 416
DB 361 ESMAYLTSGMLDQPGFPDCSIEAAMVKVFSSEAAQCVCSEALQILGCGGYMKDYPYERML 420
QY 417 RDRILLIFEGTNEILRMVIALTGLQAGRLITRIHELKQAKVSTVMDTVGRRRLDSLG 476
DB 421 RDRILLIFEGTNEILRFLIALTGLQAGRLITSRIKELKSGNVTTVMETIGRKLDSLG 480
QY 477 RTVDLGLTGNHGVHPSLADSANKPEENTCYGRVETLLRFGKTIMEEOLVLRKANI 536
DB 481 RTVDLGLTGLGVHPSLGDSANKLEENVHFGRTVETLLRFGKNIVEEQVLKRVANI 540
QY 537 LNLNGMTAVLSRASRSIRIGLRNHDHVEVLANTFCVEAYLQNLFSLSQLDKYPENLDE 596
DB 541 LNLNGMTAVLSRASRSIRIGLRNHDHVEVLANTFCVEAYLQNLFSLSQLDKYPENLDE 600
QY 597 QIKKVSQOILEKRAYICAHPLDR 619
DB 601 QIKKVSQOILEKRAYICAHPLDR 623

RESULT 4
Q8BK76
ID Q8BK76 PRELIMINARY; PRT; 625 AA.

Q8BK76;
AC 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Very-long-chain acyl-CoA dehydrogenase VLCAD homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK075984; BAC36096.1; -
SQ SEQUENCE 625 AA; 68722 MW; 29567F60B52E6FEA CRC64;

Query Match 85.7%; Score 2703; DB 11; Length 625;
Best Local Similarity 85.7%; Pred. No. 3e-182;
Matches 534; Conservative 38; Mismatches 47; Indels 4; Gaps 1;

QY 1 MSGCGLFLR-----TTAARACRGLVSTANRRLLRTSPVRAFAKELFLGKIKKKEVFPF 56
DB 1 MSGCVLLSRGATAAAAAAASRVLRFTARRRPLHTSLQSCFAKELFLGNKQGVFPF 60
QY 57 PEVSQDELNETNQLPGVPEKFFTEVDSDRKIDQEGKIPDETLEKLSLGLFGLQVPEYG 116
DB 61 PEVSQHELSEINQFVPLEKFFTEVDSDRKIDQEGKIPDETLEKLSLGLFGLQVPEYG 120
QY 117 GLGFSNTMYSLRGEIISMDGSIITVTLAAHQAIKGLIILAGTBEQAKYLPKLASGEHIA 176
DB 121 GLGLSNTMYARLGEIISLDSITVTLAAHQAIKGLIILVGNDEQAKYLPKLSGEHIA 180
QY 177 AFLCTEPASGSDAASRATLSLSEDKKHYILNGSKVWITNGGLANFTVFAKTEVVDSDG 236
DB 181 AFLCTEPASGSDAASITQTRATLSLSEDKKYFILNGSKVWITNGGLANFTVFAKTEVVDSDG 240
QY 237 SVKDKITAFIVERDFGGVTKGPKEDKLGIRGNTCEVHFENTKIPVENILGEGVGDFKVA 296
DB 241 SKTDKMTAFIVERDFGGITNGKPKEDKLGIRGNTCEVHFENTRVPVENVLGEGVGDFKVA 300
QY 297 MNLNSGRFSMGSVVAGLLKRLIEMTAETAYACTRKQFNKRLSEFGLIQEKFALMAQKAYVM 356
DB 301 MNLNSGRFSMGSAVAGMLKKLIELTAETAYACTRKQFNRLSEFGLIQEKFALMAQKAYVM 360
QY 357 ESMYTLTAGMLDQPGFPDCSIEAAMVKVFSSEAAQCVCSEALQILGGLGYTRDYPYERIL 416
DB 361 ESMAYLTSGMLDQPGFPDCSIEAAMVKVFSSEAAQCVCSEALQILGCGGYMKDYPYERML 420
QY 417 RDRILLIFEGTNEILRMVIALTGLQAGRLITRIHELKQAKVSTVMDTVGRRRLDSLG 476
DB 421 RDRILLIFEGTNEILRFLIALTGLQAGRLITSRIKELKSGNVTTVMETIGRKLDSLG 480
QY 477 RTVDLGLTGNHGVHPSLADSANKPEENTCYGRVETLLRFGKTIMEEOLVLRKANI 536
DB 481 RTVDLGLTGLGVHPSLGDSANKLEENVHFGRTVETLLRFGKNIVEEQVLKRVANI 540
QY 537 LNLNGMTAVLSRASRSIRIGLRNHDHVEVLANTFCVEAYLQNLFSLSQLDKYPENLDE 596
DB 541 LNLNGMTAVLSRASRSIRIGLRNHDHVEVLANTFCVEAYLQNLFSLSQLDKYPENLDE 600
QY 597 QIKKVSQOILEKRAYICAHPLDR 619
DB 601 QIKKVSQOILEKRAYICAHPLDR 623

RESULT 5
Q8C0B5
ID Q8C0B5 PRELIMINARY; PRT; 625 AA.

Q9H9W4; 01-MAR-2001 (TREMBlrel. 16, Created)
01-MAR-2001 (TREMBlrel. 16, Last sequence update)
01-MAR-2003 (TREMBlrel. 23, Last annotation update)
01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein FL12506.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahara K., Masuho Y., Sasaki N.;
RT "NEO human cDNA sequencing project";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK022568; BAB14104.1;
DR HSSP; Q06319; IBUC.
DR InterPro; IPR006089; Acyl-CoA_dh.
DR InterPro; IPR006090; Acyl-CoA_dh_C.
DR InterPro; IPR006091; Acyl-CoA_dh_M.
DR InterPro; IPR006092; Acyl-CoA_dh_N.
DR InterPro; IPR005829; Sug. transporter.
DR Pfam; PF00441; Acyl-CoA_dh; 1.
DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
KW Hypothetical protein.
SQ SEQUENCE 498 AA; 55083 MW; A57B7121B8FB1310 CRC64;
Query Match 80.1%; Score 2526; DB 4; Length 498;
Best Local Similarity 100.0%; Pred. No. 6,7e-170; Indels 0; Gaps 0;
Matches 498; Conservative 0; Mismatches 0;

AC Q9H9W4; 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Very-long-chain acyl-CoA dehydrogenase VICAD homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs";
RL Nature 420:563-573 (2002).
DR EMBL; AK031820; BAC27365.1;
SQ SEQUENCE 625 AA; 68694 MW; 88B0420E6A85D48A CRC64;
Query Match 85.6%; Score 2700; DB 11; Length 625;
Best Local Similarity 85.6%; Pred. No. 4.9e-182;
Matches 533; Conservative 39; Mismatches 47; Indels 4; Gaps 1;
QY 1 MSGCGFLR-----TTAARACRGLVSTANRLLRTSPVRAFAKELFLGKIKKKEVFPF 56
DB 1 MSGCVLLSRGATAAGAARASRVLRFTARRRPLHTSLQSCSAKELFLGNIEQKGVFPF 60
QY 57 PEVSODELNEINQFVPERKFFTEEDVSRKIDQSGKIPDETLEKLSGLFGLQVPEYG 116
DB 61 PEVSQHELSEINQFVPLEKFEFTTEEDVSRKIDQSGKIPDVTLEKLSGLFGLQVPEYG 120
QY 117 GLGFSNTMYSRLEIISMDGSIITVLAHQAGILKGLIAGTEEQAKYLPKLASGEHIA 176
DB 121 GLGLSNTMYARLGETIISDASITVLAHQAGILKGLIAGTEEQAKYLPKLASGEHIA 180
QY 177 AFCLETPASGDAASIRSRATLSBKKHYILNGSKVWITNGGLANITVFAKTEVVDSDG 236
DB 181 AFCLETPASGDAASIQRAVLSBKKHYILNGSKVWITNGGLANITVFAKTEVVDSDG 240
QY 237 SYVKDITAFIVERDFGGVTKGPKDLGIRGNTCEVHFENTKIPVENILGEVGDGFKVA 296
DB 241 SKTDKMTAFIVERDFGGITNGKPKDLGIRGNTCEVHFENTKIPVENILGEVGDGFKVA 300
QY 297 MNILNSGRFSMGSVAGLLKRLIEMTAETACTRQFNKRLSEFGLIOEKFALMAQKAYVM 356
DB 301 MNILNSGRFSMGSVAGMLKRLIEMTAETACTRQFNRLSEFGLIOEKFALMAQKAYVM 360
QY 357 ESMYTLTAGMLDQPGPDSCIEAAMVKVFSSEAAQCVCSEALQILGGYTRDYPYERIL 416
DB 361 ESMYTLTAGMLDQPGPDSCIEAAMVKVFSSEAAQCVCSEALQILGGYTRDYPYERIL 420
QY 417 RDRTRILLIEGTNEILRMVIALTGLQAGRLITTRIHLELKOAKVSTVMDTVGRRLRDSIG 476
DB 421 RDRTRILLIEGTNEILRMVIALTGLQAGRLITTRIHLELKOAKVSTVMDTVGRRLRDSIG 480
QY 477 RTVDLGLTGNHGVHPSLADSANKFEENTYCFGRVETLLLRFGKTIIMEEQVLKRVANI 536
DB 481 RTVDLGLTGNHGVHPSLADSANKFEENTYCFGRVETLLLRFGKTIIMEEQVLKRVANI 540
QY 537 LNLNGMTAVLSRASIRIGLRNHDHVEVLANTFCVEAYLQNLFSLSOLDKYPENILNLYGM 596
DB 541 LNLNGMTAVLSRASIRIGLRNHDHVEVLANTFCVEAYLQNLFSLSOLDKYPENILNLYGM 600
QY 597 QIKKVSQIILEKRAYICAPLDR 619
DB 601 QIKKVSQIILEKRAYICAPLDR 623
RESULT 6
Q9H9W4 PRELIMINARY; PRT; 498 AA.
ID Q9H9W4

ID	Q9V8T1	PRELIMINARY;	PRT;	655 AA.
AC	Q9V8T1;			
DT	01-MAY-2000 (TREMBlrel. 13, Created)			
DT	01-MAY-2000 (TREMBlrel. 13, Last sequence update)			
DT	01-OCT-2002 (TREMBlrel. 22, Last annotation update)			
DE	CG7461 protein.			
DE	CG7461.			
GN	Drosophila melanogaster (Fruit fly).			
OS	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BERKELEY;			
RC	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazer E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Wan K.H., Doyle C., Baxter E.H., Andrews-Pfannkoch C., Baldwin D.,			
RA	Abriil J.F., Agbayani A., An H.-J., Bayraktaroglu L., Beasley E.M.,			
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Dubin K.J., Evangelista C.E., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabriellian A.C., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hoskin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.P.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleab J.M.,			
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,			
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,			
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,			
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,			
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,			
RA	Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,			
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,			
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,			
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,			
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;			
RT	"The genome sequence of Drosophila melanogaster."			
RL	Science 287:2185-2195(2000).			
RL	EMBL: AE003796; AAF57579.1;			
DR	HSSP; Q06319; IBUC.			
DR	FlyBase; FBgn0034432; CG7461.			
DR	InterPro; IPR006089; Acyl-CoA_dh.			
DR	InterPro; IPR006090; Acyl-CoA_dh.			
DR	InterPro; IPR006091; Acyl-CoA_dh.			
DR	InterPro; IPR006092; Acyl-CoA_dh.			
DR	InterPro; IPR006093; Acyl-CoA_dh.			
DR	PFam; PF02770; Acyl-CoA_dh.M.1.			
DR	PFam; PF02771; Acyl-CoA_dh.N.1.			
DR	PROSITE; PS00072; ACYL_COA_DH.1.			
DR	PROSITE; PS00073; ACYL_COA_DH.2.1.			
DR	PROSITE; PS00073; ACYL_COA_DH.2.1.			
DR	SEQUENCE 655 AA; 71326 MW; 00916BD701E127CB CRC64;			
Query Match	41.1%; Score 1295.5; DB 5; Length 655;			
Best Local Similarity	45.1%; Pred. No. 7 8e-83;			
Matches 280; Conservative	104; Mismatches 212; Indels 25; Gaps 9;			
14 ARACRGIVV-----STANRRLLRTSPVRAFAKELFLGKIKKKEVFPFPVSQDELNEI 67				

```
RC STRAIN-Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U28943; AAA68357.1; -
DR HSSP; Q06319; 1BUC
DR WormPep; E04F6.3; CE01217.
DR InterPro; IPR006089; Acyl-CoA_dh.
DR InterPro; IPR006090; Acyl-CoA_dh_C.
DR InterPro; IPR006091; Acyl-CoA_dh_M.
DR InterPro; IPR006092; Acyl-CoA_dh_N.
DR Pfam; PF00441; Acyl-CoA_dh_1.
DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
DR Hypothetical protein.
KW SEQUENCE 613 AA; 66171 MW; C1F2883ECE34551B CRC64;

Query Match 38.4%; Score 1210.5; DB 5; Length 613;
Best Local Similarity 43.7%; Pred. No. 7e-77;
Matches 272; Conservative 109; Mismatches 205; Indels 37; Gaps 14;

QY 7 FLRTTAAACRGLVSTANRLRTSPVPFAFAKELFLGKIKKEVFPFPEVSQDELNE 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 16 FIRLSHSAKAK---DAKPKVAAYDSP--SEVMNLFKGAVDQVFPYPLNMTDEQKE 69
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 67 -INQFLGPVEKFTTEVDVRKIDQEGKIPDETLEKLKSLGLQVPEEYVGLGFSNTWY 125
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 70 TGLVMVSPLEKMLVEVDVVKNDSDIPRAVLDFAEGLTFCVLVPPLESGGNSOM 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 126 SRIGEII-SWDGSIITVLAHQAGIGLKIILACTEQKAKYLPKIASGSHIAAFCLETPA 184
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 130 ARVAEIVGAYDLGFGVVMGAHOSIGYKGLLEGTDAQOKYLPDLATGRKFAAFALTEPT 189
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 185 SGSDAAIRSRLTSDKHKHYLNGSKVWITNGGLANIEFTVFAKTEVDSDGSVKDKITA 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 190 TGSDASSVTRAEISADGRHYLVNGKIKWISNGGFADFTVFAQTPVQKQADGSKDKMSA 249
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 245 FIVERDFGVYNGKPKDKIGIRGNTCEVHFENTKIPVENILGEVDGFKVAMNINSGR 304
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 250 FIVERAFGVTSPQEKKGIRGNTTEVHFENLIPVENILGKEGEGFKVAMNINLNGR 309
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 305 FSGSVVAGLLKRLTETAEYACTRKQFNKRLSEGLTOERFALMAQKAYVMESYTYTA 364
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 310 FGIPACTGAMKHCQTKTVHDHTTRVQFGKKGQEFNGIQEKLVENIMSKLYATESTVYMLS 369
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 365 GMLDQPGPDCSIEAAVVKVFSEAAQCVSALQILGGLGYTRDPYVERILRDPRIILI 424
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 370 SNMDR-GIKYQLEAAGKVLASENAWLVCDDAIQVHGGMGFMRETGLERVLRLRIPRI 428
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 425 FEETNEILRMVYALTGLQHAGRIILTRIHELKQAKVSTVMDTVGRRRLDSLGRTVDLGLT 484
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 429 FEGANDVURLFTALTGAQHAGK-----HLAEQA--SGVGLIGLAVSRVTG-----GNT 475
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 485 G-NHG-VVHPSLADSAKFEENTYCFGRVETLLRLFGKTIMEEQVLVLRVANILINLYG 542
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 476 GSNFGQVVDASLQDSAKVLDQIALFGQTVGGLMKHKKGIIDROYEMHRVADAINIYS 535
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 543 MTAVLSRASRSIRIGLRNHDHEVLANTFCVEAYLQNLFSLQDKYAPENLDEQIKKVS 602
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 536 SAAVLSRATYAIKNKSSADPERKA-TYYVDKAMK-----SSNRRLKDGAGSVENASKVAT 591
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 603 QQILEKRAYICA-----HPLD 618
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 592 IESLAKE--VCGNGGLTLQHPVE 612
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
Q9BUX5
ID Q9BUX5
AC Q9BUX5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to NPD002 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC001817; AAH01817.1; -
DR InterPro; IPR005829; Sug_transporter.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR SEQUENCE 188 AA; 21237 MW; BDB5855A78A6DBB6 CRC64;

Query Match 30.3%; Score 954; DB 4; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.4e-59;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 434 MYIALTGLQHAGRIILTRIHELKQAKVSTVMDTVGRRRLDSLGRTVDLGLTGNHGVHPS 493
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MYIALTGLQHAGRIILTRIHELKQAKVSTVMDTVGRRRLDSLGRTVDLGLTGNHGVHPS 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 494 LADSANKFEENTYCFGRVETLLRLFGKTIMEEQVLVLRVANILINLYGMTAVLSRASRS 553
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 LADSANKFEENTYCFGRVETLLRLFGKTIMEEQVLVLRVANILINLYGMTAVLSRASRS 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 554 ITRIGLRNHDHEVLANTFCVEAYLQNLFSLQDKYAPENLDEQIKKVSQOILEKRAYIC 613
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 ITRIGLRNHDHEVLANTFCVEAYLQNLFSLQDKYAPENLDEQIKKVSQOILEKRAYIC 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 614 AHPLDRTC 621
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 AHPLDRTC 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
Q9K784
ID Q9K784
AC Q9K784;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Butyryl-CoA dehydrogenase.
GN BH3486.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RL halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001519; BAB07205.1; -
DR HSSP; Q06319; 1BUC.
DR InterPro; IPR006089; Acyl-CoA_dh.
DR InterPro; IPR006090; Acyl-CoA_dh_C.
DR InterPro; IPR006091; Acyl-CoA_dh_M.
DR InterPro; IPR006092; Acyl-CoA_dh_N.
DR Pfam; PF00441; Acyl-CoA_dh; 1.
DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00073; ACYL_COA_DH_1; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
DR Complete proteome.
KW SEQUENCE 594 AA; 65385 MW; D959657767C1B0A9 CRC64;
```

Query Match		26.0%; Score 821; DB 16; Length 594;
Best Local Similarity		33.2%; Pred. No. 2,1e-49;
Matches		207; Conservative 105; Mismatches 210; Indels 102; Gaps 13;
Qy	45	LGKIKKKEVPPFPEVSQDELNEINQFLGPVEKFT-----EVDSSKIDQEGKIPDE 96
Db	16	LDEIDADKVF-----TPEDTBEHMGKTTEFEVNVPPVLDETEHQFD-----ISRR 66
Qy	97	TLKLSLGLFLGLOVPEVGGIGFSNTMYSLRGLGELISMGSIIVTLAAHQALGKGIILA 156
Db	67	LTTQAGELGLGALVPEVGGIGLKKISSIITKEFSRAGGFSLSGAHVGGISLPIVF 126
Qy	157	GTEQAKYLPKLASGHIAAFCLTEPASGSDAASIRSRATLSDEKHHYLINGSKVMITN 216
Db	127	GNEDOKKKYLPDLATGARIAAYALTPESSGSDALGAKTTAVLNEAGTHYVLNGEKWITN 186
Qy	217	GGLANITVFPAKTEVSDSGVDKKTAFIVERDFGVNGKPKDKLGRGNSCTVEHFE 276
Db	187	SAFADVPVYAKI-----DG---EHTSAFIVKDFEGVSTGPEKKMGKIGSSTRLLILE 238
Qy	277	NTKIPVENILGEVGDGFKVAMNINLSGRFSMGSVVAGLLKRLIEMTAEAECTKQFNKRL 336
Db	239	DALVPKENLGEVGKGVHIAFINILNVRGYKLGVCIGGSKRGIELAAKYANERKQKPI 298
Qy	337	SEFGLIQEKFALMAQKAYVMESMTYLTAGMLDQ-----PGPPDCSIEA 379
Db	299	AKFTLIQEKLANMAVTTAAESSIYRTGGLFEDRLGGLSEEQQKDGREVAKAIAEYAE 358
Qy	380	AMVKVFSEAAWQCVSEALQILGLGYTRYDYPYERILDRILLIFEGTNEILRMVYALT 439
Db	359	SLNVGSEALDFVADSAVQIHGGYGFMAEYETMYRDRSINRIEFTNEINRLVPGT 418
Qy	440	GLOHA--GRI--LTTRIHELKQKYSTVMTVGRRLDSLGRITVDLGLTGNHGVHPSLAD 496
Db	419	ILRKAMGELPFLEKATALQELMMLPQEVG-----LQNLFS--LSQLDK 588
Qy	497	SANKFEENTYCFGTVTETLL-----RGKTIMEQLVLKRVANILINLYGMTAVLSR 549
Db	451	-DEPLEQEKYLLSMAKKVFIAGTGAQTYGKELQEQOELLANVADIVDSIFSMEVILR 509
Qy	550	ASRST-RIGLRNHDHVELLANTFCVEAY-----LQNLFS--LSQLDK 588
Db	510	TEKAIRNGLKBAQKALQVFCQEAFAFNTEAHAKESLVAMQSGDTRLTMTSILRKLTR 569
Qy	589	YAPENLDDQIKKVSQILEKRAYI 612
Db	570	HTPINVIAKKREIAGILEARYV 593
RESULT 13		
ID	O32176	
AC	O32176;	PRELIMINARY; PRT; 594 AA.
DT	01-JAN-1998 (TremBLrel. 05, Created)	
DT	01-JAN-1998 (TremBLrel. 05, Last sequence update)	
DT	01-MAR-2003 (TremBLrel. 23, Last annotation update)	
DE	YUSJ protein.	
GN	YUSJ.	
OS	Bacillus subtilis.	
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.	
OX	NCBI_TaxID=1423;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=168;	
RX	MEDLINE=98044033; PubMed=9384377;	
RA	Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,	
RA	Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,	
RA	Borriss R., Bourschi C., Brans A., Braun M., Brignell S.C., Bron S.,	
RA	Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,	
RA	Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,	
RA	Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,	
RA	Entian K.D., Evington J., Fabret C., Ferrari E., Foulger D.,	
RA	Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,	
RA	Ghim S.-Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,	
RA	Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.M., Henaut A.,	
RA	Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,	
RA	Joris B., Karamata D., Kasahara Y., Klier-Blanchard M., Klein C.,	
RA	Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,	
RA	Kurita K., Lapidus A., Lardinols S., Lauber J., Lazarevic V.,	
RA	Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,	
RA	Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,	
RA	Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,	
RA	Parvo V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,	
RA	Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,	
RA	Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,	
RA	Sato T., Scantian E., Schleich S., Schroeter R., Scoffone F.,	
RA	Sekiguchi J., Sekowska A., Seror S.J., Takahashi H., Takenaru K.,	
RA	Sorokin A., Tacconi E., Takagi T., Tanaka T., Terstra P., Tognoni A.,	
RA	Takeuchi M., Tamakoshi A., Vandenbol M., Vannier F., Vassarotti A.,	
RA	Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,	
RA	Viari A., Wambutt R., Wedler H., Wiedner K., Yasumoto K., Yata K.,	
RA	Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,	
RA	Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,	
RT	"The complete genome sequence of the gram-positive bacterium Bacillus	
RT	subtilis."	
RL	Nature 390:249-256(1997).	
RL	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=168;	
RA	Kunst F., Ogasawara N., Yoshikawa H., Danchin A.,	
RL	Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; Z99120; CAB5271.1;	
DR	HSSP; P26440; 11VH.	
DR	InterPro; IP006089; Acyl-CoA_dh.	
DR	InterPro; IP006090; Acyl-CoA_dh.C.	
DR	InterPro; IP006091; Acyl-CoA_dh.M.	
DR	InterPro; IP006092; Acyl-CoA_dh.N.	
DR	Pfam; PF00441; Acyl-CoA_dh; 1.	
DR	Pfam; PF02770; Acyl-CoA_dh; 1.	
DR	Pfam; PF02771; Acyl-CoA_dh.N; 1.	
DR	PROSITE; PS00072; ACYL_COA_DH_1; 1.	
DR	PROSITE; PS00073; ACYL_COA_DH_2; 1.	
KW	Complete proteome.	
SQ	SEQUENCE 594 AA; 65335 MW; F59BED1AA0D903CE CRC64;	
Query Match		25.9%; Score 815.5; DB 16; Length 594;
Best Local Similarity		33.4%; Pred. No. 5,2e-49;
Matches		201; Conservative 97; Mismatches 202; Indels 101; Gaps 12;
Qy	59	VSQDELNEINQFLGPVEKFTTEVDSRKIDQEGKIPDETLEKLSGLFLGLOVPEYGG 118
Db	45	IEQDVLPHI-----DDIENHQFESVRL-----LKKAGELGLGADVPYEGGL 88
Qy	119	GFSTMYSLRGLGELISMGDSITVTLLAAHQALGKGIILAGTEQAKYLPKLASGEHIAAF 178
Db	89	GLDKISSALITEKFSRAGSFLSYGAHVGGISLPIVFFGSEQKKYLPGLASGEKIAAY 148
Qy	179	CLTEPASGSDAASIRSRATLSDEKHHYLINGSKVMITNGLANITVFPAKTEVSDSGSV 238
Db	149	ALTEPGSGDALGAKTTAVLNEAGTHYVLNGEKWITNSAFADVPVYAKV-----DG--- 201
Qy	239	KDKITAFIVERDFGVNGKPKDKLGRGNSCTVEHNTKIPVENILGEVGDGFKVAMN 298
Db	202	-DKFSAFIVEKEFPVSTGPEKKMGKIGSSTRLLDQAEVPEKNLGLGKHVIAFN 260
Qy	299	ILNSGRFSMGSVVAGLLKRLIEMTAEAECTKQFNKRLSEFGLIOEKFALMAQKAYWES 358
Db	261	ILNIGRYKLAVGITGASKRVIELSAAYANORRQFKPTIAGLSLQTEKRIGMTASRLYAMES 320
Qy	359	MTYLTAGMLDQ-----PGPPDCSIEAAMVKVFSSEAQCVCSEALQIL 401
Db	321	SVYETVGLFEDNMNSQFTAEIDLKGRQRTAKSTAIEVAIECSLNKVFSGSTLIDYDEGVQIH 380
Qy	402	GGLGYTRYDYPYERILDRILLIFEGTNEILRMVYALTGLOHA--GRI--LTTRIHELKQKA 458
Db	381	GGYGFMAEYEVERAYRDSRINRIEFTNEINRLVPSFTFLKLGKGLPELFEKAQSLQEE 440

QY	459	KYSTVMDTVGRRLRDSGLRTVDLGLTGNHGVVHPSLADSANKFEBENTYFCGRTVETLL-	517
Db	441	LWMLPPEP-----SGVL-----EQEKYIVKQAKKIALFA	471
QY	518	-----RGKTTMEQOLVKRVANILINLYGTMATVLSRASRSIRT-GLURNHDEHLLANT	570
Db	472	AGLAAQKYGKATDREQEILVNVADIVSVYAMESAVLTKAIAAQAQKAVLYTEI	531
QY	571	FCVEAYLO-----NLSLSOLDKYAPENIDEQIKVSOQILEKRA	610
Db	532	FVOEAFNEIEAHAKESLIAMEGDSLRMLLSALRKLRTVTPKNVIOKKREAAAGIFEAK	591
QY	611	Y 611	
Db	592	Y 592	
RESULT 14			
Q8ENT4			
ID	Q8ENT4	PRELIMINARY;	PRT; 594 AA.
AC	Q8ENT4;		
DT	01-MAR-2003 (TREMBlrel. 23, Created)		
DT	01-MAR-2003 (TREMBlrel. 23, Last sequence update)		
DT	01-MAR-2003 (TREMBlrel. 23, Last annotation update)		
DE	Acetyl-CoA dehydrogenase.		
GN	OB2393.		
OS	Oceanobacillus theyeensis.		
OC	Bacteria; Firmicutes; Bacillales; Oceanobacillus.		
OX	NCBI_TaxID=182710;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=HTE831 / DSM 14371 / JCM 11309;		
RX	MEDLINE=42220767; PubMed=12235376;		
RA	Takami H., Takaki Y., Uchiyama I.;		
RT	"Genome sequence of Oceanobacillus theyeensis isolated from the theya		
RT	Ridge and its unexpected adaptive capabilities to extreme		
RT	environments.";		
RL	Nucleic Acids Res. 30:3927-3935 (2002).		
DR	EMBL: AP004601; BAC14349.1; -		
KW	Complete proteome.		
SO	SEQUENCE 594 AA; 65815 MW; E0E1E2012C026B92 CRC64;		
Query Match			
Best Local Similarity 32.3%; Score 789.5; DB 16; Length 594;			
Matches 202; Conservative 102; Mismatches 224; Indels 97; Gaps 12;			
QY	47	KIKKKEVFPPEVS-----QDELNEINQFLGPVEKFFTEVDVSRKID----QEGKIPD	95
Db	7	KIFKGGAFLVEDITGEDITTPEDFTDEHKMIAKTTDFVLGEV-VPKIDNLENHEFEHSV	65
QY	96	ETLEKLKSLGLFGLQVPEYGGFGFNTMYSRLGELISMDGSIIVTLAAHQAIGLKGIL	155
Db	66	DLLKAGDGLGADVPQYQYGLADLDKISSSLITEKFSRAGGFSVTHGAHVGIGSLPIVF	125
QY	156	ACTEQKAKYLPKLSAGHIAFCILTEPASGSDAASIRSRTLSDKKHYTLNGSKVWIT	215
Db	126	FGNDEQKENYLPKLTGELLAAAYALTEPSSGSDALGAKTAKLNDAGTHYILNGEKOWIT	185
QY	216	NGGLANITVFAKTEVDSDGSVDKDTAFIVERDFGVTKPKEDKLGRSGNTCEVHF	275
Db	186	NSAFADVFIVAKI-----DG-----EHTAFIVERDFGVTKPKEDKLGRSGNTCEVHF	237
QY	276	ENTKIPVENILGVDGDFKVMANITLNGRFSMGSVAGLLKRLIEMTAETACTRKQNK	335
Db	238	EDAEVPIENLLGKGRGHVIAFNILNNGRYKIAVGGVGGSKRALELAAYTNERKQNT	297
QY	336	LSEFGLQEKALMAQKAYVNESMTYLTAGMLDQ-----PGFPDGSIE	378
Db	298	ISSNLTKEKLTMAAGIYANESSVYRTVGLFEQMGALTDQKDKGQREVARISAEVQIE	357
QY	379	AAMVKVFSSEAAWOCVSEALQILGLGYTRDYPERILTRILLIFEGTNEILMYIAL	438
Db			

Db	358	CSMTKYMATELLDYVDEAVQMHGGYGFQMEYEVERAYRDSRINRIPEGTNEINRLVPG	417
QY	439	TGLOHA--GRI-LTRIHELKQAKVSTVMDTVGRRRLRDSGLRTVDLGLTGNHGVHPSLA	495
Db	418	TLIRKAIGLPLLOKQAQGLQELMMMPPEVG-----	450
QY	496	DSANKFEENTYFCGRTVETLL-----REFKTIMEEQVLKRVANILINLYGTMATVLS	548
Db	451	--TEALEQEYLLKNAKKWVLLGAGLAAQYKQLENEQEILNLANMVAEYFNMESAIL	508
QY	549	RASRSI-RIGLRNHDHEVLANTFCVEAY-----LQWLSLSOLD	587
Db	509	RTEKAIKSGEESKQKLLTYEYQAEPNRTEAEAKDILVTVEGDGTLRLMLSSRLKT	568
QY	588	KYAPENIDEQIKVSOQILEKRAYI	612
Db	569	RHTPVNVVKKREIALIIEEKYV	593
RESULT 15			
Q9CZC5			
ID	Q9CZC5	PRELIMINARY;	PRT; 169 AA.
AC	Q9CZC5;		
DT	01-JUN-2001 (TREMBlrel. 17, Created)		
DT	01-MAR-2003 (TREMBlrel. 23, Last sequence update)		
DT	01-MAR-2003 (TREMBlrel. 23, Last annotation update)		
DE	10, 11 days embryo whole body cDNA, RIKEN full-length enriched		
DE	library, clone:2810021C11 product:VERY-LONG-CHAIN ACYL-CoA		
DE	DEHYDROGENASE VLCD homolog (Fragment).		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Body;		
RA	Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,		
RA	Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,		
RA	Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,		
RA	Imotani K., Ishii Y., Itoh M., Izawa M., Kesukawa T., Kato H.,		
RA	Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,		
RA	Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,		
RA	Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,		
RA	Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,		
RA	Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,		
RA	Tejima Y., Toya I., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,		
RA	Muramatsu M., Hayashizaki Y.;		
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Body;		
RX	MEDLINE=22354683; PubMed=12466851;		
RA	The FANTOM Consortium,		
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;		
RT	"Analysis of the mouse transcriptome based on functional annotation of		
RT	60,770 full-length cDNAs.";		
RL	Nature 420:563-573 (2002).		
RN	[3]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Body;		
RX	MEDLINE=21085660; PubMed=11217851;		
RA	RIKEN FANTOM Consortium;		
RT	"Functional annotation of a full-length mouse cDNA collection.";		
RL	Nature 409:685-690 (2001).		
RN	[4]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Body;		
RX	MEDLINE=99279253; PubMed=10349636;		
RA	Carninci P., Hayashizaki Y.;		
RT	"High-efficiency full-length cDNA cloning.";		
RL	Meth. Enzymol. 303:19-44 (1999).		
RN	[5]		
RP	SEQUENCE FROM N.A.		

```
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakauchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanishi M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
DR EMBL; AK012773; BAB28460.2; -.
FT NON_TER 1
SQ SEQUENCE 169 AA; 19002 MW; D1B743E8E7605A76 CRC64;

Query Match 23.3%; Score 734; DB 11; Length 169;
Best Local Similarity 88.0%; Pred. No. 4.1e-44;
Matches 146; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 454 ELKQAKYSTVMDTVGRRLRDSLGRTVDGLTGNHGVVHPSLADSANKPFENTYCFGRVTE 513
DB 2 ELKSGNVTVMETIGRKLRDSLGRTVDGLTGVVHPSLGSANKLEENHYFGRTVE 61

QY 514 TLLRFCKTMEEQVLKRVANILINLYGMTAVLSRASRSIRIGLRNHDHEVLANTFCV 573
DB 62 TLLRFCKNIVEEQVLKRVANILINLYGMTAVLSRASRSIRIGLRNHDHEVLANTFCV 121

QY 574 EAYLQNFLSLSQDKYAPENLDEQIKKVSQOILEKRAYICAHPLDR 619
DB 122 EAYFQNLFLSLSQDKNAPENLDEQIKKVSQOILEKRAYICAHPLDR 167
```

Search completed: September 3, 2003, 13:30:02
Job time : 113 secs

Result	Query	Score	Match	Length	DB	ID	Description
1	3153	100.0	621	22	AA873691	Human oxidoreducta	
2	3153	100.0	621	22	AA894839	Human protein sequ	
3	3153	100.0	621	23	AA821580	Human acyl dehydro	
4	3153	100.0	628	22	AAU23008	Novel human enzyme	
5	2864	90.8	585	21	AA841800	Human ORFX ORF1564	
6	2526	80.1	498	22	AA894077	Human protein sequ	
7	1537	48.7	306	22	AAU23012	Novel human enzyme	
8	1360	43.1	655	23	ABB06992	Mouse very long ch	
9	1356.5	43.0	653	23	ABB06991	Rat very long chai	

21-JUN-2001.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being predicted and is derived by analysis of the total score distribution.

PF	07-DEC-2000; 2000WO-US33158.	
XX	16-DEC-1999; 99US-0172367.	
PR	(INCY-) INCYTE GENOMICS INC.	
XX	Yue H, Lal P, Tang YT, Hillman JL, Baughn MR, Azimzai Y, Lu DAM;	
PI	WPI: 2001-390245/41.	
XX	N-PSDB; AAH24246.	
DR	Novel human oxidoreductase protein (ORP) useful for diagnosing,	
DR	treating and preventing cell proliferative, neurological, viral,	
XX	reproductive and autoimmune/inflammatory disorders associated with	
PT	abnormal expression of ORP	
PT		
XX	Claim 1; Page 117-119; 136pp; English.	
PS		
XX	Sequences AAB73668-AAH73694 represent 27 novel human oxidoreductase	
CC	proteins, designated ORP-1 to ORP-27 respectively, and sequences	
CC	AAH24223-AAH24249 represent cDNAs encoding ORP-1 to ORP-27. Human ORP	
CC	proteins and nucleic acids are useful for diagnosing, treating or	
CC	preventing cell proliferative disorders (e.g. arteriosclerosis,	
CC	cirrhosis, psoriasis, cancers); endocrine disorders (e.g., type I or II	
CC	diabetes mellitus, diabetes insipidus, dwarfism, hirsutism, amenorrhoea,	
CC	osteoporosis); metabolic disorders (e.g., obesity, phenylketonuria,	
CC	hypercholesterolaemia); reproductive disorders (e.g., infertility,	
CC	ovulatory and menstrual cycle defects, endometriosis, polycystic	
CC	ovary disease, disruption of spermatogenesis, impotence); neurological	
CC	disorders (e.g., epilepsy, stroke, Alzheimer's disease, Huntington's	
CC	disease, Parkinson's disease, meningitis, Creutzfeldt-Jakob disease,	
CC	cerebral palsy, muscular dystrophy, mood, anxiety and schizophrenic	
CC	disorders); viral, bacterial, fungal and parasitic infections; and	
CC	autoimmune/inflammatory disorders such as acquired immunodeficiency	
CC	syndrome (AIDS), allergies, asthma, Crohn's disease, atopic dermatitis,	
CC	gout, multiple sclerosis, rheumatoid arthritis or ulcerative colitis.	
CC	Human ORP proteins and nucleotides can be used to identify compounds	
CC	which modulate their activity or expression. ORP nucleic acid sequences	
CC	may also be used for assessing the toxicity of a test compound, to detect	
CC	upstream sequences such as promoters and regulatory elements, and to	
CC	create knock out or knock in animals or transgenic animals to model	
CC	human disease. Oligonucleotide primers derived from ORP gene sequences	
CC	may be used to detect single nucleotide polymorphisms (SNPs) and for	
CC	mapping the naturally occurring genomic sequences. Antibodies specific	
CC	for ORP proteins may be used in the diagnosis of disorders associated	
CC	with aberrant ORP expression, in assays to monitor patients being treated	
CC	with ORP or modulators thereof, and for assessing toxicity of potential	
CC	drugs.	
XX		
SQ	Sequence 621 AA;	
	Query Match 100.0%; Score 3153; DB 22; Length 621;	
	Best Local Similarity 100.0%; Pred. No. 1.8e-274;	
	Matches 621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MSGCGLFRTAAARACRGLVSTANRRLRTSPVRAFAPAKELFLGKIKKKEVFPPEVS 60	
DB	1 MSGCGLFRTAAARACRGLVSTANRRLRTSPVRAFAPAKELFLGKIKKKEVFPPEVS 60	
QY	61 QDELNEINQFLGPVEKFTTEVDNRKIDQEGKIPDETLEKLSLGLFGLQVPEYGGGLG 120	
DB	61 QDELNEINQFLGPVEKFTTEVDNRKIDQEGKIPDETLEKLSLGLFGLQVPEYGGGLG 120	
QY	121 SNTMYSLRGELTISMDGSIITVLAHQAIKGLKGIILAGTEQKAKYLPKLAGSHEHIAFCL 180	
DB	121 SNTMYSLRGELTISMDGSIITVLAHQAIKGLKGIILAGTEQKAKYLPKLAGSHEHIAFCL 180	
QY	181 TEPASGSDAASIRATLSEDKHYILNGSKVWITNGGLANITVFVAKTEVWDSGVKD 240	
DB	181 TEPASGSDAASIRATLSEDKHYILNGSKVWITNGGLANITVFVAKTEVWDSGVKD 240	
QY	241 KITAFIVERDFGGVNTGKPEDKLGIRGNTCEVHFNTKIPVENILGEVGDGFKVAMNII 300	

241 KITAFIVERDFGGVNTGKPEDKLGIRGNTCEVHFNTKIPVENILGEVGDGFKVAMNII 300

301 NSGRFSGSVVAGLLKRLTETMTAEYACTRKQFNKRLSEFGLIOEKFPALMAQKAYVMESMT 360

301 NSGRFSGSVVAGLLKRLTETMTAEYACTRKQFNKRLSEFGLIOEKFPALMAQKAYVMESMT 360

361 YLTAGMLDQPGFDDCSIEAAMVKVFSSEAAWQCVSEALQILGLGTYTRDYPYERILLDR 420

361 YLTAGMLDQPGFDDCSIEAAMVKVFSSEAAWQCVSEALQILGLGTYTRDYPYERILLDR 420

421 ILLIFECTNEILRMYTALTGLOHAGRLITRIHELKQAKVSTVMDTVGRRLRDSLGR 480

421 ILLIFECTNEILRMYTALTGLOHAGRLITRIHELKQAKVSTVMDTVGRRLRDSLGR 480

481 LGLTGNHGVVHPSIADSKANKFEENTYCFGRVETLLRFCKTITMEEOLVUKRVANILINL 540

481 LGLTGNHGVVHPSIADSKANKFEENTYCFGRVETLLRFCKTITMEEOLVUKRVANILINL 540

541 YGMTAVLSRASRSIRIGLRNHDHVEVLLANTFCVEAYLONLFSLSOLDKYAPENLDEQIKK 600

541 YGMTAVLSRASRSIRIGLRNHDHVEVLLANTFCVEAYLONLFSLSOLDKYAPENLDEQIKK 600

601 VSQOILEKRAYICARPLDRTC 621

601 VSQOILEKRAYICARPLDRTC 621

RESULT 2

AAH94839

ID AAB94839 standard; Protein: 621 AA.

XX AAB94839;

XX AAB94839;

DT 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:16010.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

XX and/or diagnosis of the abnormality of the proteins encoded by the

XX full-length cDNAs -

XX Claim 8; SEQ ID 16010; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesising 5602

XX full-length cDNAs defined in the specification. Where a primer set

XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary

XX to the complementary strand of a polynucleotide which comprises one of

XX the 5602 nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95993 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 621 AA;

Query Match 100.0%; Score 3153; DB 22; Length 621;
Best Local Similarity 100.0%; Pred. No. 1.8e-274;
Matches 621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSCGCLPLRTAAARACRGLVSTANRRLLRTSPVRAFAKELFLGKIKKKEVFPFPEVS 60
Db 1 MSCGCLPLRTAAARACRGLVSTANRRLLRTSPVRAFAKELFLGKIKKKEVFPFPEVS 60
QY 61 QDELNEINQFLGPVEKFEFEVDNRKIDQEGKIPDETLEKLSGLGLQVPEYGGGLG 120
Db 61 QDELNEINQFLGPVEKFEFEVDNRKIDQEGKIPDETLEKLSGLGLQVPEYGGGLG 120
QY 121 SNTMSYRLGIIISMDGSIITVLAHQAIKGLKGIILAGTEQKAKYLPKLSGHEHIAAFCL 180
Db 121 SNTMSYRLGIIISMDGSIITVLAHQAIKGLKGIILAGTEQKAKYLPKLSGHEHIAAFCL 180
QY 181 TPASGSDAASIRSRATLSDEKHYILNGSKVITNGGLANIETVFAKTEVSDSGSKD 240
Db 181 TPASGSDAASIRSRATLSDEKHYILNGSKVITNGGLANIETVFAKTEVSDSGSKD 240
QY 241 KITAFIVERDFGGVTKGPKDKLGRSNTCEVHFNTKIPVENILGEVGDGKFAVNIL 300
Db 241 KITAFIVERDFGGVTKGPKDKLGRSNTCEVHFNTKIPVENILGEVGDGKFAVNIL 300
QY 301 NSGRFSGSVVAGLLKRLIEMTAEYACTRQFNKRLSEFGLIQEKFALMAQKAYVMSMT 360
Db 301 NSGRFSGSVVAGLLKRLIEMTAEYACTRQFNKRLSEFGLIQEKFALMAQKAYVMSMT 360
QY 361 YLTAGMLDQGFPCDSTEAAWVFSSEAAWQCVSEALQILGLGTYRDPYERILRDTR 420
Db 361 YLTAGMLDQGFPCDSTEAAWVFSSEAAWQCVSEALQILGLGTYRDPYERILRDTR 420
QY 421 ILLIFEGTNEILRMVIALTGLQHAGRLITRIHELKQAKYSTVMDTVGRRLRDSLGRTVD 480
Db 421 ILLIFEGTNEILRMVIALTGLQHAGRLITRIHELKQAKYSTVMDTVGRRLRDSLGRTVD 480
QY 481 LGLTGNHVVHPSLAUSANKFENTYCFGRVTETLLRLFRCKTMEBQVLKRVANILINL 540
Db 481 LGLTGNHVVHPSLAUSANKFENTYCFGRVTETLLRLFRCKTMEBQVLKRVANILINL 540
QY 541 YGMTAVLSRASRSIRIGLRNHDHVELLANTFCVEAYLQNLFSLSOLDKYAPENLDQIKK 600
Db 541 YGMTAVLSRASRSIRIGLRNHDHVELLANTFCVEAYLQNLFSLSOLDKYAPENLDQIKK 600
QY 601 VSOQILEKRAYICAHPLDRTC 621
Db 601 VSOQILEKRAYICAHPLDRTC 621

RESULT 3

AAE21680

ID AAE21680 standard; ProteIn; 621 AA.

XX

AC AAE21680;

XX
Dt 16-JUL-2002 (first entry)
XX
DE Human acyl dehydrogenase DHDR-7 (62112 protein).
XX
KW Human: dehydrogenase-7; DHDR-7; 62112 protein; CNS disorder;
KW epilepsy; Alzheimer's disease; AS; Pick's disease;
KW differentiation disorder; Huntington's disease; autonomic
KW function disorder; hyperthyroidism; depression; schizophrenia;
KW panic migraine; cardiac related disorder; anxiety; obesity;
KW arteriosclerosis; restenosis; Parkinson's disease; angina;
KW hypertension; cardiomyopathy; arrhythmia; muscle weakness;
KW arterial inflammation; cell proliferation disorder; growth disorder;
KW diabetes mellitus; hypertension; migration disorder; gene therapy;
KW fertility disorder; autoimmune disorder; metabolic disorder; ataxia;
KW cancer; hepatic disease; acyl-CoA dehydrogenase; beta-oxidation;
KW enzyme.
XX
OS Homo sapiens.
XX
Key Domain Location/Qualifiers
XX 1..36 /note= "Dehydrogenase domain"
XX Domain 23..123 /note= "Dehydrogenase precursor domain"
XX Domain 37..426 /note= "Acyl-coA very long chain dehydrogenase domain"
XX Domain 70..432 /note= "Dehydrogenase-related long acyl-coA chain oxidoreductase domain"
XX Domain 74..448 /note= "Short chain related acyl-coA dehydrogenase specific domain"
XX Domain 85..438 /note= "Acyl CoA dehydrogenase domain"
XX Domain 85..177 /note= "Acyl CoA dehydrogenase middle domain"
XX Domain 147..431 /note= "ACD-3 acyl-coA dehydrogenase domain"
XX Domain 172..349 /note= "Oxidoreductase acyl-coA dehydrogenase family domain"
XX Domain 179..286 /note= "Acyl CoA dehydrogenase C-terminal domain"
XX Region 179..191 /note= "Acyl CoA dehydrogenase signature 1"
XX Domain 207..604 /note= "Acyl-coA oxidase dehydrogenase oxidoreductase flavoprotein domain"
XX Domain 290..441 /note= "Acyl CoA dehydrogenase N-terminal domain"
XX Region 399..418 /note= "Acyl CoA dehydrogenase signature 2"
XX Domain 408..611 /note= "Dehydrogenase butyryl domain"
XX Domain 432..580 /note= "Polysaccharide deacetylase domain"
XX Domain 438..621 /note= "Very long chain dehydrogenase domain"
XX Region 460..476 /note= "Sugar transport protein signature"
XX
WO200218582-A2.
XX
PD 07-MAR-2002:
XX
XX 31-AUG-2001; 2001WO-US27186.
XX
PF 31-AUG-2000; 2000US-229831P.
XX
PR
XX (MILL-) MILLENNIUM PHARM INC.
PA

XX Meyers R, Hunter JJ;
PI WPI: 2002-329775/36.
DR N-PSDB; AAD34006.
XX
PT New human dehydrogenase polypeptide for diagnosing and treating
PT dehydrogenase-7 associated diseases and disorders e.g. Alzheimer's
PT disease, and to identify modulators of therapeutic use
XX
PS Claim 13; Fig 1; 124pp; English.
XX
CC The invention relates to human dehydrogenase (DHDR)-7 polypeptides
CC referred to as 62112 and nucleic acid molecules encoding such
CC polypeptides. DHDR-7 is an acyl-CoA dehydrogenase, a mitochondrial
CC flavoprotein which catalyses the first step of fatty acid beta-
CC oxidation. Sequences of the invention and their antibodies are
CC useful for treating a disorder, disease or condition which is caused
CC by misregulation (e.g. downregulation or upregulation) of DHDR activity.
CC Examples of disorders include central nervous system (CNS) disorders
CC e.g. Alzheimer's disease (AS), Pick's disease, Parkinson's, Huntington's
CC disease, Gilles de la Tourette's syndrome, multiple sclerosis, epilepsy,
CC amyotrophic lateral sclerosis and Creutzfeldt-Jakob disease; autonomic
CC function disorders e.g. hypertension, depression, schizophrenia, panic
CC migraine, anxiety, obsessive-compulsive disorder and obesity; cardiac
CC related disorder e.g. arteriosclerosis, ischaemia reperfusion injury,
CC restenosis, arterial inflammation, angina, hypertension, cardiomyopathy
CC and arrhythmia; disorders of musculus skeletal system paralysis, muscle
CC weakness e.g. ataxia; cell proliferation, growth, differentiation or
CC migration disorders e.g. cancer, diabetes mellitus, hypothyroidism or
CC hyperthyroidism; reproductive or fertility disorders; autoimmune or
CC or immune deficiency disorders; hepatic disease or dysfunction and
CC metabolic disorders. They are used for screening assays, predictive
CC medicine e.g. diagnostic assays, prognostic assays, monitoring clinical
CC trails, and pharmacogenetics. Polypeptides of the invention are used
CC to identify modulators that modulate their activity. Polynucleotides
CC of the invention are used in gene therapy. The present sequence is
CC human DHDR-7.
XX
SQ Sequence 621 AA:
Query Match 100.0%; Score 3153; DB 23; Length 621;
Best Local Similarity 100.0%; Pred. No. 1.8e-274;
Matches 621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGCGFLRTTAAARACGLVYSTANRRLLRTSPVRAFAKELFLGKIKKKEVFPPEVS 60
DB 1 MSGCGFLRTTAAARACGLVYSTANRRLLRTSPVRAFAKELFLGKIKKKEVFPPEVS 60
QY 61 QDELNEINQFLGPVEKFTTEEVDSRKIDQEGKIPDETLEKLSGLFLGQVPEEYGGIGF 120
DB 61 QDELNEINQFLGPVEKFTTEEVDSRKIDQEGKIPDETLEKLSGLFLGQVPEEYGGIGF 120
QY 121 SNTMTSRIGEIIISMGDSITVTLAAHQALGKIIILAGTEEOAKYLPKLASGEHIAAFL 180
DB 121 SNTMTSRIGEIIISMGDSITVTLAAHQALGKIIILAGTEEOAKYLPKLASGEHIAAFL 180
QY 181 TEPASGSDAASIRSRATLSEDKKKHLLNGSKVWITNGGLANFTVFAKTEVVDSDGSKVD 240
DB 181 TEPASGSDAASIRSRATLSEDKKKHLLNGSKVWITNGGLANFTVFAKTEVVDSDGSKVD 240
QY 241 KITAFIVERDFGVNGKPEDKLGIRGNSNTEVHEKTKIPVENILGEVGDGFKVAMNLT 300
DB 241 KITAFIVERDFGVNGKPEDKLGIRGNSNTEVHEKTKIPVENILGEVGDGFKVAMNLT 300
QY 301 NSGRFSMGSVVAGLLKRLIEMTAETAEYACTRKQFNKRLSEFGLTOEKFPALMAQKAYVWESMT 360
DB 301 NSGRFSMGSVVAGLLKRLIEMTAETAEYACTRKQFNKRLSEFGLTOEKFPALMAQKAYVWESMT 360
QY 361 YLTAGMLDQPGPDGSIEMAMVKVFSSEAAWQCVSEALQILGGLGYTRDYPYERILRDR 420
DB 361 YLTAGMLDQPGPDGSIEMAMVKVFSSEAAWQCVSEALQILGGLGYTRDYPYERILRDR 420

QY 421 ILLIFEGTNEILRMVIALTGLQHAGRIILTRIHHELKQAKVSTVMDTVGRRRLDSIGRTVD 480
DB 421 ILLIFEGTNEILRMVIALTGLQHAGRIILTRIHHELKQAKVSTVMDTVGRRRLDSIGRTVD 480
QY 481 LGLTGNHGVVHPSLADSSANKFEENTYCFGRVETILLRFGKTIMEEQLVLRVANILNL 540
DB 481 LGLTGNHGVVHPSLADSSANKFEENTYCFGRVETILLRFGKTIMEEQLVLRVANILNL 540
QY 541 YGMTAVLSRASRSIRIGLRNHDHVEYLLANTFCVEAYLQNLFSLSQDKYAPENLDEQIKK 600
DB 541 YGMTAVLSRASRSIRIGLRNHDHVEYLLANTFCVEAYLQNLFSLSQDKYAPENLDEQIKK 600
QY 601 VSQQLERKRAYICAHPLDRTC 621
DB 601 VSQQLERKRAYICAHPLDRTC 621
RESULT 4
AAU23008
ID AAU23008 standard; Protein; 628 AA.
XX
AC AAU23008;
XX
DT 17-DEC-2001 (first entry)
XX
DE Novel human enzyme polypeptide #94.
XX
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
KW nephrotropic; anticoagulant.
XX
OS Homo sapiens.
XX
PN WO200155301-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01239.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-022547.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229513.
PR 05-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-02321968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241788.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.

PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2000US-0254097.
PR 05-JAN-2001; 2000US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI: 2001-465566/50.

N-PSDB: AAS40878.

Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases

Claim 11: SEQ ID No 1004; 1180pp; English.

The present invention relates to the isolation of novel human enzyme polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of a wide range of disorders, including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS), autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. influenza). The polynucleotides of the invention can also be used in gene therapy. AAU22915-AAU23814 represent the novel human enzyme polypeptides of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

Sequence 628 AA;

Query Match 100.0%; Score 3153; DB 22; Length 628;
Best Local Similarity 100.0%; Pred. No. 1.8e-274;
Matches 621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Tue Sep 9 10:14:37 2003

us-09-945-326-2.rag

```
QY 1 MSGCGFLRTTAAARACGLVSTANRLLRTSPVRAFAKELFLGKTKKEVPPPEVS 60
Db 8 MSGCGFLRTTAAARACGLVSTANRLLRTSPVRAFAKELFLGKTKKEVPPPEVS 67
QY 61 QDELNEINOFGLPVKEFFTEEDVSRKIDQEGKIPDETLEKLSGLFLGQVPEYGGIGF 120
Db 68 QDELNEINOFGLPVKEFFTEEDVSRKIDQEGKIPDETLEKLSGLFLGQVPEYGGIGF 127
QY 121 SNTMYSRGELIISMDGSIITVTLAAHQAIKGLGIIILAGTEEOAKYLPKLASGEHIAFCL 180
Db 128 SNTMYSRGELIISMDGSIITVTLAAHQAIKGLGIIILAGTEEOAKYLPKLASGEHIAFCL 187
QY 181 TEPASGSDAASIRSRATLSEDKKHVILNGSKVWIITNGGLANFTVFAKTEVVDSGSKVD 240
Db 188 TEPASGSDAASIRSRATLSEDKKHVILNGSKVWIITNGGLANFTVFAKTEVVDSGSKVD 247
QY 241 KITAFIVERDFGGVTKGPKEDKLGIRGNTCEVHFENTKIPVENILGEVGDGFKVAMNIL 300
Db 248 KITAFIVERDFGGVTKGPKEDKLGIRGNTCEVHFENTKIPVENILGEVGDGFKVAMNIL 307
QY 301 NSGRFSMGSVAGLLKRLIEMTAETKQFNKRLSEFGLIOEKFALMAQKAYVME 360
Db 308 NSGRFSMGSVAGLLKRLIEMTAETKQFNKRLSEFGLIOEKFALMAQKAYVME 367
QY 361 YLTAGMLDQPGPDSCIEAAMVKVFSSEAAWQCVSEALQILGGLGYTRDYPYERILRDTR 420
Db 368 YLTAGMLDQPGPDSCIEAAMVKVFSSEAAWQCVSEALQILGGLGYTRDYPYERILRDTR 427
QY 421 ILLIFGVTNEILRMVYALTGLQAGHRLITRHELKQAKVSVMTVGBRLDSLGRTVD 480
Db 428 ILLIFGVTNEILRMVYALTGLQAGHRLITRHELKQAKVSVMTVGBRLDSLGRTVD 487
QY 481 LGTGNHGVVHPSLADSKANKFENTYCFGRVETLLRFGKTIEMEQVLKRVANILNL 540
Db 488 LGTGNHGVVHPSLADSKANKFENTYCFGRVETLLRFGKTIEMEQVLKRVANILNL 547
QY 541 YGMTAVLSRASRSIRIGLRNHDHEVLLANTFCVEAYLQNLFSLSQDKYAPENLDEQIKK 600
Db 548 YGMTAVLSRASRSIRIGLRNHDHEVLLANTFCVEAYLQNLFSLSQDKYAPENLDEQIKK 607
QY 601 VSQIILEKRAYICAHPLDRFC 621
Db 608 VSQIILEKRAYICAHPLDRFC 628
```

RESULT 5

AB41800

ID AB41800 standard; Protein; 565 AA.

XX AC AAB41800;

XX DT 08-FEB-2001 (first entry)

XX DE Human ORFX ORF1564 polypeptide sequence SEQ ID NO:3128.

Human; open reading frame; ORFX; detection; cytostatic; hepatotropic; vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac; immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antiviral; antibacterial; antifungal; antirheumatic; antithyroid; antianaemic; gene therapy; cancer; proliferative disorder; hypertension; neurodegenerative disorder; osteoarthritis; graft vs host disease; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; cholesterol ester storage; systemic lupus erythematosus; infection; severe combined immunodeficiency; malaria; autoimmune disorder; asthma; allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antiinflammatory disease; coagulation; thrombosis; contraceptive.

OS Homo sapiens.

XX WO200058473-A2.

PN

```
XX 05-OCT-2000.
XX 31-MAR-2000; 2000WO-US086621.
XX 31-MAR-1999; 99US-0127607.
XX 02-APR-1999; 99US-0127636.
XX 05-APR-1999; 99US-0127728.
XX 30-MAR-2000; 2000US-0540763.
XX (CURA-) CURAGEN CORP.
XX Shimkets RA, Leach M;
XX WPI; 2000-602362/57.
XX N-PSDB; AAC76009.
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
XX Claim 11; Page 2345-2346; 5507pp; English.
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnary;
XX antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
XX osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
XX immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; antirheumatic;
XX antiinflammatory; antibacterial; antiviral; antifungal; antianaemic;
XX antithyroid; and antianaemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
XX Sequence 565 AA;
XX Query Match 90.8%; Score 2864; DB 21; Length 565;
XX Best Local Similarity 100.0%; Pred. No. 1.5e-248; Indels 0; Gaps 0;
XX Matches 564; Conservative 0; Mismatches 0;
QY 58 EVSQDELNEINOFGLPVKEFFTEEDVSRKIDQEGKIPDETLEKLSGLFLGQVPEYGG 117
Db 2 EVSQDELNEINOFGLPVKEFFTEEDVSRKIDQEGKIPDETLEKLSGLFLGQVPEYGG 61
QY 118 LGFSNTMYSRGELIISMDGSIITVTLAAHQAIKGLGIIILAGTEEOAKYLPKLASGEHIAA 177
Db 62 LGFSNTMYSRGELIISMDGSIITVTLAAHQAIKGLGIIILAGTEEOAKYLPKLASGEHIAA 121
QY 178 FCUTEPASGSDAASIRSRATLSEDKKHVILNGSKVWIITNGGLANFTVFAKTEVVDSG 237
Db 122 FCUTEPASGSDAASIRSRATLSEDKKHVILNGSKVWIITNGGLANFTVFAKTEVVDSG 181
QY 238 VKDKITAFIVERDFGGVTKGPKEDKLGIRGNTCEVHFENTKIPVENILGEVGDGFKVAM 297
Db 182 VKDKITAFIVERDFGGVTKGPKEDKLGIRGNTCEVHFENTKIPVENILGEVGDGFKVAM 241
QY 298 NILNSGRFSMGSVAGLLKRLIEMTAETKQFNKRLSEFGLIOEKFALMAQKAYVME 357
Db 242 NILNSGRFSMGSVAGLLKRLIEMTAETKQFNKRLSEFGLIOEKFALMAQKAYVME 301
QY 358 SMTYLTAGMLDQPGPDSCIEAAMVKVFSSEAAWQCVSEALQILGGLGYTRDYPYERILR 417
Db 358 SMTYLTAGMLDQPGPDSCIEAAMVKVFSSEAAWQCVSEALQILGGLGYTRDYPYERILR 417
```

Db 302 SMTYLTAGMLDQPGFPCDSEIAAMVKVFSSEAAHQVCSEALQILGGIGYTRDYPYRILR 361
 QY 418 DTRILLIFEGTNEILRMVIALTGLQHAGRILTTIRIHELKQAKVSTVMDTVGRRRLDSLGR 477
 Db 362 DTRILLIFEGTNEILRMVIALTGLQHAGRILTTIRIHELKQAKVSTVMDTVGRRRLDSLGR 421
 QY 478 TVDLGLTGNHGVVHPSLADSSANKPEENTYCFGRVTETLLRFRGKTIMEEOLVLRVANIL 537
 Db 422 TVDLGLTGNHGVVHPSLADSSANKPEENTYCFGRVTETLLRFRGKTIMEEOLVLRVANIL 481
 QY 538 INLYGMPAVLSRASRSIRIGLRNHDHVEVLANTFCVEAYLQNLFSQLDKYAPENLDEQ 597
 Db 482 INLYGMPAVLSRASRSIRIGLRNHDHVEVLANTFCVEAYLQNLFSQLDKYAPENLDEQ 541
 QY 598 IKKVSQOILEKRAYICAHPLDRTC 621
 Db 542 IKKVSQOILEKRAYICAHPLDRTC 565

RESULT 6

AAB94077
 ID AAB94077 standard; Protein: 498 AA.
 XX AC AAB94077;
 XX DT 26-JUN-2001 (first entry)
 XX DE Human protein sequence SEQ ID NO:14271.
 XX KW Human: primer; detection; diagnosis; antisense therapy; gene therapy.
 XX OS Homo sapiens.
 XX PN EP1074617-A2.
 XX PD 07-FEB-2001.
 XX PF 28-JUL-2000; 2000EP-0116126.
 XX PR 29-JUL-1999; 99JP-0248036.
 XX PR 27-AUG-1999; 99JP-0300253.
 XX PR 11-JAN-2000; 2000JP-0118776.
 XX PR 02-MAY-2000; 2000JP-0183767.
 XX PR 09-JUN-2000; 2000JP-0241899.
 XX PA (HELI-) HELIX RES INST.
 XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

Claim 8; SEQ ID 14271; 2537pp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB93893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

XX SQ Sequence 498 AA;

Query Match 80.1%; Score 2526; DB 22; Length 498;

Best Local Similarity 100.0%; Pred. No. 3.2e-218;

Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 MYSRLGEIISMDGSITVTAAHQAIGLKGIILAGTEEQAKYLPKLASGEHIAAFCLTEP 183
 Db 1 MYSRLGEIISMDGSITVTAAHQAIGLKGIILAGTEEQAKYLPKLASGEHIAAFCLTEP 60
 QY 184 ASGSDAASIRSRATLSSEDKKHYILNGSKVITNGLANIFTVFAKTEVVDSDGSKDKIT 243
 Db 61 ASGSDAASIRSRATLSSEDKKHYILNGSKVITNGLANIFTVFAKTEVVDSDGSKDKIT 120
 QY 244 AFIVERDEGGVTNGKPEDKLGIRSNTECEVHFNTKIPVENILGEVGDGFKVAMNINSG 303
 Db 121 AFIVERDEGGVTNGKPEDKLGIRSNTECEVHFNTKIPVENILGEVGDGFKVAMNINSG 180
 QY 304 RFSMGVVAGLLKRLIEMTAETACTRQFNKRLSEFGLIOEKFPALMAQKAYVMSWYLT 363
 Db 181 RFSMGVVAGLLKRLIEMTAETACTRQFNKRLSEFGLIOEKFPALMAQKAYVMSWYLT 240
 QY 364 AGMLDQGFPCDSEIAAMVKVFSSEAAHQVCSEALQILGGIGYTRDYPYRILRDRILL 423
 Db 241 AGMLDQGFPCDSEIAAMVKVFSSEAAHQVCSEALQILGGIGYTRDYPYRILRDRILL 300
 QY 424 IFECTNEILRMVIALTGLQHAGRILTTIRIHELKQAKVSTVMDTVGRRRLDSLGRYVDLGL 483
 Db 301 IFECTNEILRMVIALTGLQHAGRILTTIRIHELKQAKVSTVMDTVGRRRLDSLGRYVDLGL 360
 QY 484 TGNHGVVHPSLADSSANKPEENTYCFGRVTETLLRFRGKTIMEEOLVLRVANILINLYGM 543
 Db 361 TGNHGVVHPSLADSSANKPEENTYCFGRVTETLLRFRGKTIMEEOLVLRVANILINLYGM 420
 QY 544 TAVLSRASRSIRIGLRNHDHVEVLANTFCVEAYLQNLFSQLDKYAPENLDEQIKKVSQ 603
 Db 421 TAVLSRASRSIRIGLRNHDHVEVLANTFCVEAYLQNLFSQLDKYAPENLDEQIKKVSQ 480
 QY 604 QILEKRAYICAHPLDRTC 621
 Db 481 QILEKRAYICAHPLDRTC 498

RESULT 7

AAU23012

ID AAU23012 standard; Protein: 306 AA.

XX AC AAU23012;

XX AC AAU23012;

DT 18-DEC-2001 (first entry)

XX DE Novel human enzyme polypeptide #98.

KW Human: oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
 KW ligase; hyperproliferative disorder; immunodeficiency disorder;
 KW autoimmune disorder; neurological disorder; metabolic disorder;
 KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
 KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
 KW nephrotropic; anticoagulant.

XX OS Homo sapiens.

XX XX

PN WO200155301-A2.

XX

Tue Sep 9 10:14:37 2003

us-09-945-326-2.rag

PD	02-AUG-2001.	PR	29-SEP-2000;	2000US-0236368.
XX		PR	29-SEP-2000;	2000US-0236369.
PF	17-JAN-2001;	PR	29-SEP-2000;	2000US-0236370.
XX	2001WO-US01239.	PR	02-OCT-2000;	2000US-0236802.
XX		PR	02-OCT-2000;	2000US-0237037.
PR	31-JAN-2000;	PR	02-OCT-2000;	2000US-0237038.
PR	04-FEB-2000;	PR	02-OCT-2000;	2000US-0237039.
PR	24-FEB-2000;	PR	02-OCT-2000;	2000US-0237040.
PR	02-MAR-2000;	PR	13-OCT-2000;	2000US-0239935.
PR	16-MAR-2000;	PR	13-OCT-2000;	2000US-0239937.
PR	17-MAR-2000;	PR	20-OCT-2000;	2000US-0240960.
PR	18-APR-2000;	PR	20-OCT-2000;	2000US-0241221.
PR	19-MAY-2000;	PR	20-OCT-2000;	2000US-0241785.
PR	07-JUN-2000;	PR	20-OCT-2000;	2000US-0241786.
PR	28-JUN-2000;	PR	20-OCT-2000;	2000US-0241787.
PR	30-JUN-2000;	PR	20-OCT-2000;	2000US-0241808.
PR	07-JUL-2000;	PR	20-OCT-2000;	2000US-0241809.
PR	07-JUL-2000;	PR	20-OCT-2000;	2000US-0241826.
PR	11-JUL-2000;	PR	01-NOV-2000;	2000US-0244617.
PR	11-JUL-2000;	PR	08-NOV-2000;	2000US-0246474.
PR	14-JUL-2000;	PR	08-NOV-2000;	2000US-0246475.
PR	26-JUL-2000;	PR	08-NOV-2000;	2000US-0246476.
PR	26-JUL-2000;	PR	08-NOV-2000;	2000US-0246477.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246478.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246523.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246524.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246525.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246526.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246527.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246528.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246532.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246609.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246610.
PR	14-AUG-2000;	PR	08-NOV-2000;	2000US-0246611.
PR	18-AUG-2000;	PR	08-NOV-2000;	2000US-0246613.
PR	22-AUG-2000;	PR	17-NOV-2000;	2000US-0249207.
PR	22-AUG-2000;	PR	17-NOV-2000;	2000US-0249208.
PR	22-AUG-2000;	PR	17-NOV-2000;	2000US-0249209.
PR	22-AUG-2000;	PR	17-NOV-2000;	2000US-0249210.
PR	30-AUG-2000;	PR	17-NOV-2000;	2000US-0249211.
PR	01-SEP-2000;	PR	17-NOV-2000;	2000US-0249212.
PR	01-SEP-2000;	PR	17-NOV-2000;	2000US-0249213.
PR	01-SEP-2000;	PR	17-NOV-2000;	2000US-0249214.
PR	01-SEP-2000;	PR	17-NOV-2000;	2000US-0249215.
PR	05-SEP-2000;	PR	17-NOV-2000;	2000US-0249216.
PR	05-SEP-2000;	PR	17-NOV-2000;	2000US-0249217.
PR	06-SEP-2000;	PR	17-NOV-2000;	2000US-0249218.
PR	06-SEP-2000;	PR	17-NOV-2000;	2000US-0249244.
PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249245.
PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249264.
PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249265.
PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249297.
PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249299.
PR	08-SEP-2000;	PR	17-NOV-2000;	2000US-0249300.
PR	12-SEP-2000;	PR	01-DEC-2000;	2000US-0250160.
PR	12-SEP-2000;	PR	01-DEC-2000;	2000US-0250391.
PR	14-SEP-2000;	PR	05-DEC-2000;	2000US-0251030.
PR	14-SEP-2000;	PR	05-DEC-2000;	2000US-0251988.
PR	14-SEP-2000;	PR	05-DEC-2000;	2000US-0256719.
PR	14-SEP-2000;	PR	06-DEC-2000;	2000US-0251479.
PR	14-SEP-2000;	PR	08-DEC-2000;	2000US-0251856.
PR	14-SEP-2000;	PR	08-DEC-2000;	2000US-0251868.
PR	14-SEP-2000;	PR	08-DEC-2000;	2000US-0251869.
PR	21-SEP-2000;	PR	08-DEC-2000;	2000US-0251989.
PR	21-SEP-2000;	PR	11-DEC-2000;	2000US-0251990.
PR	25-SEP-2000;	PR	05-JAN-2001;	2001US-0254097.
PR	25-SEP-2000;	PR		2001US-0259678.
PR	26-SEP-2000;	XX		
PR	27-SEP-2000;	PA	(HUMA-) HUMAN GENOME SCI INC.	
PR	27-SEP-2000;	XX	Rosen CA, Barash SC, Ruben SM;	
PR	29-SEP-2000;	PI		
PR	29-SEP-2000;	XX	WPI; 2001-465566/50.	
PR	29-SEP-2000;	XX		
PR	29-SEP-2000;	DR		

DR N-PSDB; AAS40882.

XX Novel polypeptides and polynucleotides useful for diagnosing,
 PT preventing, treating neural, immune system, muscular, reproductive,
 PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous
 PT diseases

XX
 PS Claim 11: SEQ ID No 1008; 1180pp; English.

XX
 CC The present invention relates to the isolation of novel human enzyme
 CC polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences
 CC encoding them. The enzyme polypeptides of the invention may comprise the
 CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
 CC isomerases or ligases. The sequences of the invention are useful in the
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of
 CC disorders, including hyperproliferative disorders (e.g. cancer),
 CC immunodeficiency disorders (e.g. AIDS), autoimmune disorders
 CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
 CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders
 CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
 CC blood-related disorders (e.g. haemophilia), reproductive disorders
 CC (e.g. infertility) and infectious disorders (e.g. Influenza). The
 CC polynucleotides of the invention can also be used in gene therapy.
 CC AAU22915-AAU23814 represent the novel human enzyme polypeptides of the
 CC invention.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 306 AA;

Query Match 48.7%; Score 1537; DB 22; Length 306;
 Best Local Similarity 99.3%; Pred. No. 1.4e-129;
 Matches 302; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 318 LIETAEYACTRKQFNKRLSEGLIOEKFAALMAOKAYVWESMTYLTAGMLDQGFDDCSI 377
 DB 3 LXEMTAETACTRKQFNKRLSEGLIOEKFAALMAOKAYVWESMTYLTAGMLDQGFDDCSI 62

QY 378 EAMVKVFSSEAAWOCVSEALQILGGLGYTRDYPYERILRDRILLIPEGTNEILRMVIA 437
 DB 63 EAMVKVFSSEAAWOCVSEALQILGGLGYTRDYPYERILRDRILLIPEGTNEILRMVIA 122

QY 438 LTGLQHAGRIITLRIHELKQAKVSTVMDTVGRRRLDLSGRVVDLGLGNHGVHPHSLADS 497
 DB 123 LTGLQHAGRIITLRIHELKQAKVSTVMDTVGRRRLDLSGRVVDLGLGNHGVHPHSLADS 182

QY 498 ANKFEENTYCFGRVTETLLRFGKTIIMEEQVLKRVANILINLGMTAVLSRSRSTRIG 557
 DB 183 ANKFEENTYCFGRVTETLLRFGKTIIMEEQVLKRVANILINLGMTAVLSRSRSTRIG 242

QY 558 LRNHDEVLANTFCVAYLQNLFLSLQDKYAPENLDEQIKKYSQOILEKRAYICAHPL 617
 DB 243 LRNHDEVLANTFCVAYLQNLFLSLQDKYAPENLDEQIKKYSQOILEKRAYICAHPL 302

QY 618 DRTC 621
 DB 303 DRTC 306

RESULT 8
 ABB06992
 ID ABB06992 standard; Protein; 655 AA.

XX
 AC ABB06992;
 XX
 DT 20-JUN-2002 (first entry)

XX Mouse very long chain acyl-CoA dehydrogenase (VLCAD) protein SEQ ID:10.
 DE Human; MD25; VLCAD; very long chain acyl-CoA dehydrogenase; IRAP;
 KW Insulin responsive aminopeptidase; GLUT4; glucose transporter 4;
 KW antidiabetic; intracellular regulation; glucose metabolism; diabetes;
 KW

KW glucose transport; hyperglycaemic disorder.
 XX Mus sp.
 XX WO200216428-A1.
 XX 28-FEB-2002.
 XX
 PF 20-AUG-2001; 2001WO-JP071117.
 XX
 PR 21-AUG-2000; 2000JP-0254263.
 PR 07-SEP-2000; 2000JP-0276633.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Tojo H, Katayama N, Kakimoto S;
 XX WPI; 2002-269344/31.
 DR
 XX Protein binding to insulin-responsive aminopeptidase and glucose
 PT transporter 4, useful for prevention and treatment of diseases
 PT associated with blood sugar level disturbance
 XX
 PS Disclosure; Page 92-95; 103pp; Japanese.

XX The present invention describes human MD25 (very long chain acyl-CoA
 CC dehydrogenase (VLCAD)), which binds to insulin-responsive aminopeptidase
 CC (IRAP) and to glucose transporter 4 (GLUT4). MD25 has antidiabetic
 CC activity. IRAP and GLUT4 are involved in the intracellular regulation of
 CC glucose metabolism and glucose transport across the cell membrane.
 CC Expression of MD25 (whose ligands are these proteins) is also involved
 CC in this regulatory process. MD25 can be used in the prevention, treatment
 CC and diagnosis of diseases involving disturbances of glucose metabolism,
 CC such as diabetes and other hyperglycaemic disorders. The present sequence
 CC represents mouse VLCAD which is given in the exemplification of the
 CC present invention.

XX Sequence 655 AA;

Query Match 43.1%; Score 1360; DB 23; Length 655;
 Best Local Similarity 48.2%; Pred. No. 3.9e-113;
 Matches 287; Conservative 106; Mismatches 184; Indels 18; Gaps 9;

QY 35 PVRA-FAKELFLGKIKKKEVFPPEV-SODELNEINFLGPVEKFEFEVDSRKIDQ 89
 DB 66 PARAESKSPAVGFKGOLTIQVFPSPVLSSEQAQFLKELVGPVAFVEEVNDPAKND 125

QY 90 EGKIPDETLEKLSLGLFGLQVPEEYGGGLFSNTMYSLRGLIISM-DGSITVTTLAAHAI 148
 DB 126 LEKVEDDTLQGLKELGAFGLQVPSLGLGLSLSTQVRLAEIVGMHDLGVSVTLGAHQSI 185

QY 149 GLKGIILAGTEEQAKYLPKLASGEHIAAFCLTEPASGSDAASIRSRATLSDEKHHYILN 208
 DB 186 GFGLIILYGTAKREYLPKRVASGQALAAFCLETPSSGSDVASIRSAIPSPCGKYITLN 245

QY 209 GSKWITNGGLANIFTVFARTEVVD-SDGSVKDKITAFIVRDFGVTKGPKDKIGRG 267
 DB 246 GSKWISNGGLADIFTVFARTEVVD-SDGSVKDKITAFIVRDFGVTKGPKDKIGRG 305

QY 268 SNTCEVHFENTKIPVENILCEVGDGPKVAMNINLSGRFSNGSVAGLLKRLIEMTAAYAC 327
 DB 306 SNTSEVYFDGVKVPSENVILGEVGDGPKVAMNINLSGRFSNGSVAGLLKRLIEMTAAYAC 365

QY 328 TRKQFNKRLSEGLIOEKFAALMAOKAYVWESMTYLTAGMLDQGFDDCSI 387
 DB 366 NRTQFGCDKTHNFGVIOEKFAALMAOKAYVWESMTYLTAGMLDQGFDDCSI 424

QY 388 EAMWOCVSEALQILGGLGYTRDYPYERILRDRILLIPEGTNEILRMVIA 447
 DB 425 EAMWOCVSEALQILGGLGYTRDYPYERILRDRILLIPEGTNEILRMVIA 484

QY 448 LTRIHELKQ--AKYSTVMDTVGRRRLDLSGRVVDLGLGNHGVHPHSLADSANKFENT 505

Db	485	LTGLGNALKNPFNVGLLMGAGKQLRRRTGIGSGLSL	----GIVHPELSRSGELAYQAL	541
Qy	506	YCFGRVETLLRFRGKTIMEQVLKRVANILINLYGMTAVLSRASRSIRIGLRNHHDEV	565	
Db	542	DQFATVVEAKLVKHKKGVNEQFLQRLADGAILDYAMVVVLSRASRSLSGEGYPTAQHEK	601	
Qy	566	LLANTFCVEA---YLQNLFSLSQLDKYAPENLDEQIKKVSQOILEKRAYICAHPL	617	
Db	602	MLCDSWCIEAATRIRENMASSQSPQH--QELFRNFRSISKAMVNGGLVTCNPL	654	
RESULT 9				
ID	ABB06991	standard; Protein; 653 AA.		
XX	AC	ABB06991;		
XX	DT	20-JUN-2002 (first entry)		
XX	DE	Rat very long chain acyl-CoA dehydrogenase (VLCAD) protein SEQ ID NO:9.		
XX	KW	Human; MD25; VLCAD; very long chain acyl-CoA dehydrogenase; IRAP;		
XX	KW	insulin responsive aminopeptidase; GLUT4; glucose transporter 4;		
XX	KW	antidiabetic; intracellular regulation; glucose metabolism; diabetes;		
XX	KW	glucose transport; hyperglycaemic disorder.		
XX	OS	Rattus sp.		
XX	PN	WO200216428-A1.		
XX	PD	28-FEB-2002.		
XX	PF	20-AUG-2001; 2001WO-JP07117.		
XX	PR	21-AUG-2000; 2000JP-0254263.		
XX	PR	07-SEP-2000; 2000JP-0276633.		
XX	PA	(TAKE) TAKEDA CHEM IND LTD.		
XX	PI	Tojo H, Katayama N, Kakimoto S;		
XX	DR	WPI; 2002-269344/31.		
XX	PT	Protein binding to insulin-responsive aminopeptidase and glucose		
XX	PT	transporter 4, useful for prevention and treatment of diseases		
XX	PS	associated with blood sugar level disturbance		
XX	PS	Disclosure; Page 90-92; 103pp; Japanese.		
XX	CC	The present invention describes human MD25 (very long chain acyl-CoA		
XX	CC	dehydrogenase (VLCAD)), which binds to insulin-responsive aminopeptidase		
XX	CC	(IRAP) and to glucose transporter 4 (GLUT4). MD25 has antidiabetic		
XX	CC	activity. IRAP and GLUT4 are involved in the intracellular regulation of		
XX	CC	glucose metabolism and glucose transport across the cell membrane.		
XX	CC	Expression of MD25 (whose ligands are these proteins) is also involved		
XX	CC	in this regulatory process. MD25 can be used in the prevention, treatment		
XX	CC	and diagnosis of diseases involving disturbances of glucose metabolism,		
XX	CC	such as diabetes and other hyperglycaemic disorders. The present sequence		
XX	CC	represents rat VLCAD which is given in the exemplification of the		
XX	CC	present invention.		
XX	SQ	Sequence 653 AA;		
Query Match				
Best Local Similarity 43.08; Score 1356.5; DB 23; Length 653;				
Matches 295; Conservative 108; Mismatches 192; Indels 41; Gaps 12;				
Qy	9	RTTAARACGLVVSTANRLLR-----TSPVRA-----FAKELFLGKIKKKEV	53	
Db	32	RPTSAQR-----LVASEATQVLEKPELSSDASTREKPARAESKSFVAGNFKQLTDOV	87	
Qy	54	FPPEVSSQDELNE-----INQFLGPVEKFTTEEVDSRKIDQEGKIPDETLEKLSGLFG	108	
109 LOVPEEYGGIGFSNTYSRLGEIISM-DGSITVTTLAAHQAIQGLKGIILLAGTEEQAKYLP				
144 LOVSELGGLGLSNTQVARLAEIVGMIDLGVSYTLGAHQSIGFGKILLYCTKAQKEKYL				
168 KLASGEHIAAFCLTEPASGSDAASIRSRATLSSEDKHYIILNGSKVWITNGSLANIFVFA				
204 RVASGQALAAFLTEPSSGSDVARSIRSAVSPGCKYTYTLNGSKIMWISNGGLADIFTVFA				
228 KTEVVD-SDGSVKDKITAFIVERDFGVTKNGKPKDKLGRSNTCEVHFENTKIPVENIL				
264 KTPIKDAATGAVKEKITAFVVERSGVTHGLPEKKMGIKASNTSEVYFDGVKVPANVL				
287 GEVGDGFKVAMNINLSGRFSMGVYVAGLLKRLIEMTAETAYACTRKQFNKRLSEFGLIOEKF				
324 GEVGDGFKVAVNINLNGRFGMAATLAGTMTKAIIAKAVDHATNRTQFGDKIHNFVGIOEKL				
347 ALMAOKAYVMESWYLTAGMLDQPGFPCDTEAAWVKVFSSEAAWOCVSEALQILGLGY				
384 ARMAILQYVTESMAYMLSANMDQ-GFKDFQTEAAISKIFGSEAAWCVTDECIQIMGMGF				
407 TRDYPYERILDRTRILLIFECTNELRMVIALTGILQHAGRIITLTRIHELKO--AKVSTVM				
443 MKEPGEVERLDIRIFRIFECTNDILRLFVALQCMCKELTGLGNALKNPLGNVGLLI				
465 DTGVRRLRDSLGRTVDLGTGNHGVHPVSLADSANKFEENTYCFGRVETILLRFGKTIM				
503 GEASKQLRRRTGIGSGLSL---GIVHPELSRSGELAVOALEQFATVVEAKLMHKKGIV				
525 EEQLVKKEVANILINLYGMTAVLSRASRSIRIGLRNHDHVLANTFCVEA---YLQNL				
560 NEQFLQLRDLADGAILDYAMVVVLSRASRSLSGEGYPTAQHEKMLCDSWCIEAATRIENMA				
582 SILSQLDKYAPENLDEQIKKVSQOILEKRAYICAHPL				
620 SILQSNPQ--QQELFRNFRSISKAMVNGGLVTSNPL				
RESULT 10				
ID	ABB06993	standard; Protein; 655 AA.		
XX	AC	ABB06993;		
XX	DT	20-JUN-2002 (first entry)		
XX	DE	Bovine very long chain acyl-CoA dehydrogenase (VLCAD) protein SEQ ID:11.		
XX	KW	Human; MD25; VLCAD; very long chain acyl-CoA dehydrogenase; IRAP;		
XX	KW	insulin responsive aminopeptidase; GLUT4; glucose transporter 4;		
XX	KW	antidiabetic; intracellular regulation; glucose metabolism; diabetes;		
XX	OS	glucose transport; hyperglycaemic disorder.		
XX	OS	Bos taurus.		
XX	PN	WO200216428-A1.		
XX	PD	28-FEB-2002.		
XX	PF	20-AUG-2001; 2001WO-JP07117.		
XX	PR	21-AUG-2000; 2000JP-0254263.		
XX	PR	07-SEP-2000; 2000JP-0276633.		
XX	PA	(TAKE) TAKEDA CHEM IND LTD.		
XX	PI	Tojo H, Katayama N, Kakimoto S;		
XX	DR	WPI; 2002-269344/31.		
XX	PT	Protein binding to insulin-responsive aminopeptidase and glucose		
XX	PT	transporter 4, useful for prevention and treatment of diseases		

	QY	394	VSEALQILGGLYGTRDYPYERILDRDTRILLIFESTNEILRMVIALTGLQHAGKRLITRIH	453
Dd		430	TDECIOIMGGMGFKPEGVVERVLRLRFRIFEFNSTNDILFLFVALQCMDKKELSGLS	489
	QY	454	ELUK--AKVSVMTDVGRRLRDSLGRTVDLGLTGNGHVHPHSADSANKFEENTYCFCGT	511
Dd		490	ALKNPFGNAGLLGLEAGQLRRRAGLSGSLS---GLVHPESURSGEIAVRALEOFATV	546
	QY	512	VETLLLRFCKTIMEQLVKLVANILNLXGMTAVLSRASRSIRIGLRNHHDHEVLANTF	571
Dd		547	VBAKLTKHKGVINEQFFLRLADCAIDLAYAMVVVLSRASRSLSGEHTPAHQEKMLCDTW	606
	QY	572	CVE--AYLNQLFSOLDKYAPENIDEQIKKVSOOILEKRATCAHPL	617
Dd		607	CJFAAARIREGMAALOSDPWOEE-LYRNFKSIGSKALVERGVVTISNP	653

RESULT 12

ABB63264
ID ABB63264 standard: Protein: 655 AA.

XX ABB63264;

XX
DT 26-MAR-2002 (first entry)

XX
DE
proscribila melanogaster polypeptide SEO ID NO 16584.

XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX
05
prosophila melanoqaster.

XX
PN
WO200171042-A2.

XX
PD
27-SEP-2001.

XX 23-MAR-2001: 2001WO-US09231.
PF

XX
PR 23-MAR-2000: 2000US-191637P.

PR 23-MAR-2000; 2000US-191037F.
PR 11-JUL-2000; 2000US-0614150.

XX
PA (PEKE) PE CORP NY.

XX
PI
Venter JC, Adams M, Li PWD, Myers EW;

XX
DB WPT: 2001-656860/75.

DR WPI; 2001-636880/
DR N-PSDB: ABL07367.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
PT interactions -

XX
PS
disclosure: SEO ID NO 16584: 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL10176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

CC (ABB57737-ABB/2072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published pct sequences.

XX	Sequence	655 AA;
50		

Query Match 41.1%; Score 1295.5; DB 22; Length 655;
Best Local Similarity 45.1%; Pred. No. 2.5e-107;
Matches 280; Conservative 104; Mismatches 212; Indels 25; Gaps 9;

14 A B A C B C I V W - - - - - S T A N R R I . I . R T S P P V R A F A K E L F L G K I K K K E V F F P P E V S Q D E L N E I 67

Db	44	ASLCRQATATHSPKLGAESNRSKEKASEN--ESFAMNFRGSLVSSVFFPDPVLTAEOKEK	102
QY	68	-NQPLGPEVKEFFTEEDVDSRKIDQEGKIPDETLEKLSLGLFGLQVPEEVGGGLGFSNTMYS	126
Db	103	TNSLIDPFERFSDVNDAAHRANDANSKIDDTSTTALMELGAFGIQVSEFGGLGLNNTQYG	162
QY	127	RLGETISM-DGSIIVTTLAAHQATLGLKGIILAGTPEQAKYLPKLKASGHEHIAAFCLTEPAS	185
Db	163	RLCALVGVNDLGLGITTCAGHSIGFKGILLYGTPPEOKERYLPKVAEEQVYAAAFALTEPSS	222
QY	186	GSDAASIRSRATLSDEKKHYILNGSKVWITNGGLANFIYFAKTEVVD--SDGSVKDKLTA	244
Db	223	GSDAGSIRCAVKSADGKHVYVNLNGSKTIWISNGGIAEIMTVFAQTQEQVDPKTEKKDKVTA	282
QY	245	FIVERDFCGVTNGKPEDKIGTIGSNTCEVHFENTKIPVENILGEVGDGFKVAMNINLSGR	304
Db	283	FIVERSFGGVNGPEPKWKGIKASNTAEVYFEDVKIPNIENVLGKESGDFKVAAMNINLSGR	342
QY	305	FSWGSVAVAGLLRLRIEMTAEVACTRKOFNKRLSSEFGLIQEKFPALMAQAAYVMESWTYLT	364
Db	343	FCMGATLSGTMKKCTEQATEHANNRVQFGKLNKYSIQEKLQAMNIOYATESMAFTIS	402
QY	365	GMLDQPGPDSCSIEAAMVKFSEEAACQVSEALQILGLGLGYTRDYPYERILRDLTRILLI	424
Db	403	QNND--AGSKDHYLEAAISKIYVASEAWTVCDVAIQILGGMGYMVDNGLERVLDRIFRI	461
QY	425	FEGTNEILRMVYALTGLQHAGRIILTRIHLEKQ-----AKVSTVMDTVGRRRLDSLGR	478
Db	462	FEGTNDILRLFTALTGQYAG---SHLKLQRAFKNPSANLGLIFKEASR-----AAS	513
QY	479	VDLGLTGNHGVHPSLADSANKFEENTYFCGRWTETILLRFGKTIIMEQVLVKRVAANILI	538
Db	514	VGLGGTDLGSHGVVGGELLYPAKTAHCIDLFGSGVEEILLRYKNKINVEQIILLRLANA	573
QY	539	NYLGMTAVLSRASRSTRIGLRNHDHVELLANTFCVEAYLQNLSOLDKAPENLDEQI	598
Db	574	DIYAMVVTQSRSSRAVNLNLPQAQHELNLTALTOASDRVIKNLQAAATSSHRSLSNEKI	633
QY	599	KKVSQOQILEKRAYICAHPLDR	619
Db	634	STIAKTTLENGGVTTTGILDQ	654
RESULT 13			
ID	AAU23009	standard; Protein; 202 AA.	
XX	AAU23009;		
XX	17-DEC-2001	(first entry)	
XX	Nucleol human enzyme polypeptide #95.		

RESULT 13

RESULT 13
AA123009

AAU23009
ID AAU23009 standard; Protein; 202 AA.

AA
AC
AAU23009;

17-DEC-2001 (first entry)

XX
DE
Nervol human enzyme polypeptide #95.

XX Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
 KW ligase; hyperproliferative disorder; immunodeficiency disorder;
 KW autoimmune disorder; neurological disorder; metabolic disorder;
 KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
 KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
 KW neurotropic; anticoagulant.

XX Homo sapiens.

XX WO200155301-A2. DN

XX 02-AUG-2001.

XX
PE 17-JAN-2001: 2001WO-US01239.

XX
PP
31 -JAN-2000: 2000US-0179065.
XX

PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225466.
PR 14-AUG-2000; 2000US-0225467.
PR 14-AUG-2000; 2000US-0225467.
PR 14-AUG-2000; 2000US-0225468.
PR 14-AUG-2000; 2000US-0225468.
PR 14-AUG-2000; 2000US-0225477.
PR 14-AUG-2000; 2000US-0225477.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465566/50.

N-PSDB; AAS40879.

Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases

Claim 11; SEQ ID No 1005; 1180pp; English.

DR XX The present invention relates to the isolation of novel human enzyme
 CC polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences
 CC encoding them. The enzyme polypeptides of the invention may comprise the
 CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
 CC isomerases or ligases. The sequences of the invention are useful in the
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of
 CC disorders including hyperproliferative disorders (e.g. cancer),
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders
 CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
 CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders
 CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
 CC blood-related disorders (e.g. haemophilia), reproductive disorders
 CC (e.g. infertility) and infectious disorders (e.g. Influenza). The
 CC polynucleotides of the invention can also be used in gene therapy.
 CC AAU22915-AAU23814 represent the novel human enzyme polypeptides of the
 CC invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 202 AA;
 Query Match 31.1%; Score 980; DB 22; Length 202;
 Best Local Similarity 100.0%; Pred. No. 9.7e-80;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSGCGLFLRTTAAARACRGLVSTANRLLRTSPVRAFAKELFLGKIKKEVFPPEVS 60
 DB 8 MSGCGLFLRTTAAARACRGLVSTANRLLRTSPVRAFAKELFLGKIKKEVFPPEVS 67
 QY 61 QDELNEINQFLGPVEKFFTEEVDSRKIDQEGKIPDETLEKLSGLGLQVPEEYGGIGF 120
 DB 68 QDELNEINQFLGPVEKFFTEEVDSRKIDQEGKIPDETLEKLSGLGLQVPEEYGGIGF 127
 QY 121 SNTMYSRLGEIISMDGSIITVTLAAHQAIIGLKGIIAGTTEOKAKYLPKL 180
 DB 128 SNTMYSRLGEIISMDGSIITVTLAAHQAIIGLKGIIAGTTEOKAKYLPKL 187
 QY 181 TEPASGSDAASIRSR 195
 DB 188 TEPASGSDAASIRSR 202
 RESULT 14
 ABP76260
 ID ABP76260 standard; Protein: 169 AA.
 XX
 AC ABP76260;
 XX
 DT 21-FEB-2003 (first entry)
 XX
 DE Human GENSET protein SEQ ID 810.
 XX
 KW Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;
 KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
 KW inflammatory disease; immune disorder; neuromuscular; toxicity;
 KW central nervous system; cardiovascular; gastrointestinal.
 XX
 OS Homo sapiens.
 XX
 PN WO20028398-A1.
 XX
 PD 24-OCT-2002.
 XX
 PF 18-APR-2001; 2001WO-IB00914.
 XX
 PR 18-APR-2001; 2001WO-IB00914.
 XX
 PA (GEST) GENSET.
 XX
 PI Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;
 XX

WPI: 2003-075548/07.
 New GENSET polynucleotides and polypeptides, useful for treating heavy
 metal toxicity, cancer, inflammatory diseases, immune disorders, and
 the neuromuscular, CNS, cardiovascular or gastrointestinal effects of
 the toxicity
 Claim 14; Page 690; 735pp; English.
 The present invention relates to novel GENSET polynucleotides
 (ABZ36404-ABZ36911) encoding polypeptides (ABP75963-ABP76368). The
 polynucleotides and polypeptides are useful in screening and diagnostic
 assays for abnormal GENSET expression and/or biological activity. They
 are also useful for screening of compounds for treating or preventing
 GENSET-related disorders, such as heavy metal toxicity, cancer, central
 inflammatory diseases, immune disorders, and the neuromuscular, central
 nervous system (CNS), cardiovascular or gastrointestinal effects of the
 toxicity.
 Sequence 169 AA;
 Query Match 24.5%; Score 774; DB 24; Length 169;
 Best Local Similarity 91.7%; Pred. No. 2.5e-61;
 Matches 155; Conservative 2; Mismatches 12; Indels 0; Gaps 0;
 QY 1 MSGCGLFLRTTAAARACRGLVSTANRLLRTSPVRAFAKELFLGKIKKEVFPPEVS 60
 DB 1 MSGCGLFLRTTAAARACRGLVSTANRLLRTSPVRAFAKELFLGKIKKEVFPPEVS 60
 QY 61 QDELNEINQFLGPVEKFFTEEVDSRKIDQEGKIPDETLEKLSGLGLQVPEEYGGIGF 120
 DB 61 QDELNEINQFLGPVEKFFTEEVDSRKIDQEGKIPDETLEKLSGLGLQVPEEYGGIGF 120
 QY 121 SNTMYSRLGEIISMDGSIITVTLAAHQAIIGLKGIIAGTTEOKAKYLPKL 169
 DB 121 SNTMYSRLGEIISMDGSIITVTLAAHQAIIGLKGIIAGTTEOKAKYLPKL 169
 RESULT 15
 ABB60259
 ID ABB60259 standard; Protein: 414 AA.
 XX
 AC ABB60259;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 7569.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2001; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB; ABL04362.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX

Search completed: September 3, 2003, 13:27:31
Job time : 87 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2003, 10:14:14 ; Search time 513.782 seconds
(without alignments)
9788.298 Million cell updates/sec

Title: US-09-945-326-3
Perfect score: 1863
Sequence: 1 atagcggctgcgggtctt.....acccctggacagacatgc 1863

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_19Jun03:*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	1863	100.0	2452	24	AAD34006	Human acyl dehydro
2	1863	100.0	2522	23	ABV21123	Human prostate exp
3	1863	100.0	2522	23	ABV23498	Human prostate exp
4	1863	100.0	2522	23	ABV26965	Human prostate exp
5	1863	100.0	2522	23	ABV29355	Human prostate exp
6	1862.6	100.0	2492	22	AAS40878	cDNA encoding nove
7	1861.4	99.9	2440	22	AAH24246	Human oxidoreducta
8	1859.8	99.8	2399	22	AAH16781	Human cDNA sequenc

9	1590.8	90.8	2218	21	AAC76009	Human ORFX ORF1564
10	1671.4	89.7	2400	22	AAH15804	Human CDNA sequenc
11	912.6	49.0	1118	22	AAS40882	cDNA encoding nove
12	704.6	37.8	771	22	AAH08631	Human CDNA clone (
13	585	31.4	651	22	AAS40879	cDNA encoding nove
14	499.6	26.8	772	20	AAZ158901	Human validated ca
15	490.4	26.3	835	20	AAZ15890	Human gene express
16	484.4	26.0	543	25	ABZ36713	Human GENSET codin
17	460.4	24.7	507	25	ABZ36843	Human GENSET codin
18	418.4	22.5	431	22	AAF67278	Novel human polynu
19	393.2	21.1	796	22	AAH05329	Human CDNA clone (
20	391.6	21.0	1965	24	ABL50829	Human MD25 (VLCAD)
21	391.6	21.0	2148	24	ABL50834	Human MD25 (VLCAD)
22	388.6	20.9	2117	24	ABK63677	Rat sequence diffe
23	382.8	20.4	434	22	AAF65775	Novel human polynu
24	372.8	20.0	2045	23	ABL12894	Drosophila melanog
25	372.8	20.0	4045	23	ABL12894	Drosophila melanog
26	372.6	20.0	412	22	AAF65875	Novel human polynu
27	358.4	19.2	2186	21	AAC93369	Human secreted pro
28	318.8	17.1	321	22	AAS41476	cDNA encoding nove
29	300	16.1	300	20	AAZ13131	Human gene express
30	283.2	15.2	300	20	AAZ13131	Human cancer cell
31	259.2	13.9	472	23	ABV44576	Human prostate exp
32	233.8	12.5	252	23	ABV14697	Human prostate exp
33	225.4	12.1	275	23	ABV35775	Human prostate exp
34	221.8	11.9	1024	25	ABZ83763	Toxicologically re
35	214.6	11.5	312	24	ABK35342	Human CDNA encodin
36	197.8	10.6	462	23	ABV32112	Human prostate exp
37	196.4	10.5	458	23	ABV41049	Human prostate exp
38	191.8	10.3	221	23	ABV01794	Human prostate exp
39	183.4	9.8	185	23	ABV10963	Human prostate exp
40	159.4	8.6	183	24	ABL80414	Human ovarian canc
41	146.6	7.9	241	20	AAX39501	Gene #2309 used to
42	146.6	7.9	1829	24	ABN95811	Mycobacterium tube
43	143.4	7.7	4403765	22	AAI99683	Mycobacterium tube
44	143.4	7.7	4411529	22	AAI99682	Drosophila melanog
45	140.4	7.5	1569	23	ABL06885	

ALIGNMENTS

RESULT 1

AAD34006

ID AAD34006 standard; CDNA; 2452 BP.

XX AAD34006;

AC AAD34006;

XX AAD34006;

DT 16-JUL-2002 (first entry)

XX Human acyl dehydrogenase DHDR-7 (62112 protein)-encoding CDNA.

DE Human; dehydrogenase-7; DHDR-7; 62112 protein; CNS disorder;

XX Human; dehydrogenase-7; DHDR-7; 62112 protein; CNS disorder;

KW epilepsy; Alzheimer's disease; AS; Pick's disease;

KW differentiation disorder; Huntington's disease; autonomic

KW function disorder; hyperthyroidism; depression; schizophrenia;

KW panic migraine; cardiac related disorder; anxiety; obesity;

KW arteriosclerosis; restenosis; Parkinson's disease; angina;

KW hypertension; cardiomyopathy; arrhythmia; muscle weakness;

KW arterial inflammation; cell proliferation disorder; growth disorder;

KW diabetes mellitus; hypertension; migration disorder; gene therapy;

KW fertility disorder; autoimmune disorder; metabolic disorder; ataxia;

KW cancer; hepatic disease; acyl-CoA dehydrogenase; beta-oxidation;

KW gene; ss.

XX Homo sapiens.

OS Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

FT	3'UTR	/note="This region is specifically claimed in claim 1 as SEQ ID NO:3"
FT	FT	1933..2452
FT	FT	/*tag= C
FT	FT	
FT	FT	
PN	WO200219582-A2.	
XX	XX	
PD	07-MAR-2002.	
PF	31-AUG-2001; 2001WO-US27186.	
PR	31-AUG-2000; 2000US-229831P.	
XX	(MILL-) MILLENNIUM PHARM INC.	
PA	Meyers R, Hunter JJ;	
PI	WPI: 2002-329775/36.	
DR	P-PSDB; AAR21680.	
XX	New human dehydrogenase polypeptide for diagnosing and treating dehydrogenase-7 associated diseases and disorders e.g. Alzheimer's disease, and to identify modulators of therapeutic use -	
PT	PT	
PT	PT	
XX	XX	Claim 1; Fig 1; 124pp; English.
XX	XX	The invention relates to human dehydrogenase (DHDR)-7 polypeptides referred to as 62H12 and nucleic acid molecules encoding such polypeptides. 62H12 is an acyl-CoA dehydrogenase, a mitochondrial flavoprotein which catalyses the first step of fatty acid beta-oxidation. Sequences of the invention and their antibodies are useful for treating a disorder, disease or condition which is caused by misregulation (e.g. downregulation or upregulation) of DHDR activity. Examples of disorders include central nervous system (CNS) disorders e.g. Alzheimer's disease (AS), Pick's disease, Parkinson's, Huntington's disease, Gilles de la Tourette's syndrome, multiple sclerosis, epilepsy, any trophic lateral sclerosis and Creutzfeldt-Jakob disease; autonomic function disorders e.g. hypertension, depression, schizophrenia, panic migraine, anxiety, obsessive-compulsive disorder and obesity; cardiac related disorder e.g. arteriosclerosis, ischaemia reperfusion injury, restenosis, arterial inflammation, angina, hypertension, cardiomyopathy and arrhythmia; disorders of musculus skeletal system paralysis, muscle weakness e.g. ataxia; cell proliferation, growth, differentiation or migration disorders e.g. cancer, diabetes mellitus, hypothyroidism or hyperthyroidism; reproductive or fertility disorders; autoimmune or immune deficiency disorders; hepatic disease or dysfunction and metabolic disorders. They are used for screening assays, predictive medicine e.g. diagnostic assays, prognostic assays, monitoring clinical trails, and pharmacogenetics. Polypeptides of the invention are used to identify modulators that modulate their activity. Polynucleotides of the invention are used in gene therapy. The present sequence is cDNA encoding human DHDR-7:
XX	XX	Sequence 2452 BP; 616 A; 616 C; 685 G; 535 T; 0 other;
XX	XX	
QY	Query Match	100.0%; Score 1863; DB 24; Length 2452;
Db	Best Local Similarity	100.0%; Pred. No. 0;
QY	Matches 1863; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Db	1	ATGAGCGGCTCGGGGCTTCTCTCGGCACACACGCGTCTGGGCTCTGGCTCTCGGGGGTCTG 60
QY	67	ATGAGCGGCTCGGGGCTTCTCTCGGCACACACGCGTCTGGGCTCTGGCTCTCGGGGGTCTG 126
Db	61	GTGGTCTCTACCGGAAACCGCGGCTACTTGGCGACACGCGGCTGTACGAGCTTTCGCC 120
QY	127	GTGGTCTCTACCGGAAACCGCGGCTACTTGGCGACACGCGGCTGTACGAGCTTTCGCC 186
Db	121	AAAGAGCTTTTCTTAGGCAAAATCAAGAAGAAGATTTTCCCATTTCCAGAAGTTAGC 180
QY	187	AAAGAGCTTTTCTTAGGCAAAATCAAGAAGAAGATTTTCCCATTTCCAGAAGTTAGC 246
Db	181	CAAGATGAACCTAATGAATCAATCAGTCTTCTTGGACCCGCTGGAAAAATTTCTTCACGTAA 240

QY 1321 CTGCAGCATGCGCGCCGATCCCTGACTACAGGATCCATGAGCTTAACAGGCCAAAGTG 1380
 DB 1387 CTGCAGCATGCGCGCCGATCCCTGACTACAGGATCCATGAGCTTAACAGGCCAAAGTG 1446
 QY 1381 AGCACATCATGATACCTGTTGGCGGAGGCTTCGGGACTCCCTGGGCGGAACTGTGGAC 1440
 DB 1447 AGCACATCATGATACCTGTTGGCGGAGGCTTCGGGACTCCCTGGGCGGAACTGTGGAC 1506
 QY 1441 CTGGGGCTGACAGCAACCATGAGTTGTGCACCCCGAGTCTTCGGGACAGTGCCACAAG 1500
 DB 1507 CTGGGGCTGACAGCAACCATGAGTTGTGCACCCCGAGTCTTCGGGACAGTGCCACAAG 1566
 QY 1501 TTTGAGGAGAACACCTACTGCTTCGGCGGACCGTGGAGACACTGCTGCTCCGCTTTGGC 1560
 DB 1567 TTTGAGGAGAACACCTACTGCTTCGGCGGACCGTGGAGACACTGCTGCTCCGCTTTGGC 1626
 QY 1561 AAGACCATCATGAGGAGGAGCTGTTGTAAGCGGGTGGCCAACTCCTCATCAACCTG 1620
 DB 1627 AAGACCATCATGAGGAGGAGCTGTTGTAAGCGGGTGGCCAACTCCTCATCAACCTG 1686
 QY 1621 TATGGCATGACGCGCGTGTGTCGGCGGACCGCTCCATCGGCTTGGGCTCCGCAAC 1680
 DB 1687 TATGGCATGACGCGCGTGTGTCGGCGGACCGCTCCATCGGCTTGGGCTCCGCAAC 1746
 QY 1681 CAGCACAGAGTTCTTGGGCAACACCTTCTGCGTGAAGCTTACTTGCAGATCTC 1740
 DB 1747 CAGCACAGAGTTCTTGGGCAACACCTTCTGCGTGAAGCTTACTTGCAGATCTC 1806
 QY 1741 TTCAGCTCTCTCAGCTGACAGTATGCTCCAGAAACCTAGATGAGCAGATTAAGAAA 1800
 DB 1807 TTCAGCTCTCTCAGCTGACAGTATGCTCCAGAAACCTAGATGAGCAGATTAAGAAA 1866
 QY 1801 GTGTCCAGCATCTTGGAGAGCGAGCTATATCTGTGCGCACCTCTGACAGGACA 1860
 DB 1867 GTGTCCAGCATCTTGGAGAGCGAGCTATATCTGTGCGCACCTCTGACAGGACA 1926
 QY 1861 TGC 1863
 DB 1927 TGC 1929

RESULT 2
 ID ABV21123 standard; cDNA: 2522 BP.
 XX
 AC ABV21123;
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 21114.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;

XX

DR

XX

PT

PT

PT

XX

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

WPI: 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -

Claim 1; Page 3496; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:

(a) assessing whether a patient is afflicted with prostate cancer;
 (b) monitoring the progression of prostate cancer in a patient;
 (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;
 (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;
 (e) selecting a composition for inhibiting prostate cancer in a patient;
 (f) assessing the prostate cell carcinogenic potential of a compound;
 (g) determining whether prostate cancer has metastasized in a patient;
 (h) assessing the aggressiveness or indolence of prostate cancer in a patient;

(I) is also useful as a pharmacodynamic or pharmacogenomic marker.

Sequence 2522 BP; 642 A; 634 C; 701 G; 543 T; 2 other;

Query Match

Best Local Similarity 100.0%; Score 1863; DB 23; Length 2522;

Matches 1863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCGCTCGGGCTCTTCTCGGACACGAGCTGGGCTCGTGGCTCGCGGGTCTG 60

DB 117 ATGAGCGCTCGGGCTCTTCTCGGACACGAGCTGGGCTCGTGGCTCGCGGGTCTG 176

QY 61 GTGTCTCTACCGGCAACCGGCGCTACTGCGCACGAGCGCTGTACGAGCTTCGCC 120

DB 177 GTGTCTCTACCGGCAACCGGCGCTACTGCGCACGAGCGCTGTACGAGCTTCGCC 236

QY 121 AAGAGCTTTTCCTAGGCAAAATCAAGAAGAAAGTTTCCATTTCCAGAGTTAGC 180

DB 237 AAGAGCTTTTCCTAGGCAAAATCAAGAAGAAAGTTTCCATTTCCAGAGTTAGC 296

QY 181 CAAGATGAATTAATGAATCAATCAGTCTCTGGGACCCGTTGGAATAATTTCCACATGAA 240

DB 297 CAAGATGAATTAATGAATCAATCAGTCTCTGGGACCCGTTGGAATAATTTCCACATGAA 356

QY 241 GAGTGGACTCCCGAAAAATTTGACAGGAGGAAAAATCCAGATGAATTTGGAGAAA 300

DB 357 GAGTGGACTCCCGAAAAATTTGACAGGAGGAAAAATCCAGATGAATTTGGAGAAA 416

QY 301 TTGAAGAGCTTAGGCTTTTGGCTGCAAGTCCAGAGAAATATGTTGGCTTGGCTTC 360

DB 417 TTGAAGAGCTTAGGCTTTTGGCTGCAAGTCCAGAGAAATATGTTGGCTTGGCTTC 476

QY 361 TCCAACACCTACTACTCAAGACTAGGGAGATCATCAGATGGTGGTCCATCAGTGTG 420

DB 477 TCCAACACCTACTACTCAAGACTAGGGAGATCATCAGATGGTGGTCCATCAGTGTG 536

QY 421 ACCCTGCGACCGCACCGGCTATTGGCTCAAGGGATCATTTGGCTGGCACTGAGGAG 480

DB 537 ACCCTGCGACCGCACCGGCTATTGGCTCAAGGGATCATTTGGCTGGCACTGAGGAG 596

QY 481 CAGAAAGCCAAATACTTGCCTAACTGGCTCCGGGAGCACATTCAGGCTTCTGCCCTC 540

DB 597 CAGAAAGCCAAATACTTGCCTAACTGGCTCCGGGAGCACATTCAGGCTTCTGCCCTC 656

QY 541 ACGAGCCAGCAGTGGGAGCGATGAGCTTCAATCCGGAGCAGAGCCACACTAAGTGAA 600

DB 657 ACGAGCCAGCAGTGGGAGCGATGAGCTTCAATCCGGAGCAGAGCCACACTAAGTGAA 716

QY 601 GACAAGACACTACATCTCTCAATGCTCAAGTCTGGATTACTAATGGAGACTGGCC 660

DB 601 GACAAGACACTACATCTCTCAATGCTCAAGTCTGGATTACTAATGGAGACTGGCC 660

Db 717 GACAAGAGCACTACATCTCAATGGCTCCAAGGCTCGGATTACTAATGGAGGACTGGCC 776
Qy 661 AATATTTTACTGTGTTTTCGAAGACTGAGTGGTGTGATCTCTGATGGATCACTGAAGAC 720
Db 777 AATATTTTACTGTGTTTTCGAAGACTGAGTGGTGTGATCTCTGATGGATCACTGAAGAC 836
Qy 721 AATATCACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAACCCGAA 780
Db 837 AATATCACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAACCCGAA 896
Qy 781 GATAAATAGGCAATTCGGGCTCCAAACACTTTGGAAGTCCATTTTGAAGAACCAAGATA 840
Db 897 GATAAATAGGCAATTCGGGCTCCAAACACTTTGGAAGTCCATTTTGAAGAACCAAGATA 956
Qy 841 CTTGTGGAAGAACATCTTGGAGAGTCCGAGAGTGGGTTTAAGGTGGCCATGAATCCCTC 900
Db 957 CTTGTGGAAGAACATCTTGGAGAGTCCGAGAGTGGGTTTAAGGTGGCCATGAATCCCTC 1016
Qy 901 AACAGGGCCGGTTTCAGCATGGGCTGAGCGTCTGGCTGGCTGCTCAAGAGATTGATGAA 960
Db 1017 AACAGGGCCGGTTTCAGCATGGGCTGAGCGTCTGGCTGGCTGCTCAAGAGATTGATGAA 1076
Qy 961 ATGACTGCTGAGTACGCTCCACAGGAACAGTTTAAACAGAGGCTCAGTGAATTGGA 1020
Db 1077 ATGACTGCTGAGTACGCTCCACAGGAACAGTTTAAACAGAGGCTCAGTGAATTGGA 1136
Qy 1021 TTGATTCAGGAGAAATTTGACATGATGGCTCAGAAGGCTTACGTCATGGAGAGTATGACC 1080
Db 1137 TTGATTCAGGAGAAATTTGACATGATGGCTCAGAAGGCTTACGTCATGGAGAGTATGACC 1196
Qy 1081 TACCTCACAGAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGAGCC 1140
Db 1197 TACCTCACAGAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGAGCC 1256
Qy 1141 ATGCTGAAGTGTTCACCTCGAGGCGCTGCGAGTGTGAGTGGAGGCTGCGAGATC 1200
Db 1257 ATGCTGAAGTGTTCACCTCGAGGCGCTGCGAGTGTGAGTGGAGGCTGCGAGATC 1316
Qy 1201 CTCGGGGCTTGGCTACACAAGGAGTATCCGTACAGAGGCTACGTGCTGACACCCG 1260
Db 1317 CTCGGGGCTTGGCTACACAAGGAGTATCCGTACAGAGGCTACGTGCTGACACCCG 1376
Qy 1261 ATCTCTCTCATCTTCGAGGAGCAACATGAGATTCCTCGGATGTACATCGCCCTCACGGT 1320
Db 1377 ATCTCTCTCATCTTCGAGGAGCAACATGAGATTCCTCGGATGTACATCGCCCTCACGGT 1436
Qy 1321 CTCGAGCATGCCGCCGATCCTGACTACAGGATCCATGAGCTTAAACAGGCAAAAGTG 1380
Db 1437 CTCGAGCATGCCGCCGATCCTGACTACAGGATCCATGAGCTTAAACAGGCAAAAGTG 1496
Qy 1381 AGCAGATCATGATACCGTTGGCCGGAGGCTTCGGGACTCCCTGGGCGCAACTGTGGAC 1440
Db 1497 AGCAGATCATGATACCGTTGGCCGGAGGCTTCGGGACTCCCTGGGCGCAACTGTGGAC 1556
Qy 1441 CTGGGGCTGACAGCAACCATGGAGTTGTGACACCCAGTCTTGGGAGAGTGGCAACAAG 1500
Db 1557 CTGGGGCTGACAGCAACCATGGAGTTGTGACACCCAGTCTTGGGAGAGTGGCAACAAG 1616
Qy 1501 TTTGAGGAGAACACTTACTGCTTGGCCGGACCGTGGAGACACTGCTCCCGTTTGGC 1560
Db 1617 TTTGAGGAGAACACTTACTGCTTGGCCGGACCGTGGAGACACTGCTCCCGTTTGGC 1676
Qy 1561 AAGACCATCATGGAGGAGCAGTGGTACTGAAGCGGTGGCCACATCCATCAACCTG 1620
Db 1677 AAGACCATCATGGAGGAGCAGTGGTACTGAAGCGGTGGCCACATCCATCAACCTG 1736
Qy 1621 TATGCAATGAGCGGCTGCTGCGGGCCAGCGCTCATCGCATTTGGGCTCCGCAAC 1680
Db 1737 TATGCAATGAGCGGCTGCTGCGGGCCAGCGCTCATCGCATTTGGGCTCCGCAAC 1796
Qy 1681 CACGACCAAGAGGTTCTCTTGGCCAAACACCTTCTGCTGGAGGCTTACTTCGAGATCTC 1740
Db 1797 CACGACCAAGAGGTTCTCTTGGCCAAACACCTTCTGCTGGAGGCTTACTTCGAGATCTC 1856

Qy 1741 TTCAGCCCTCTCTACGCTGGACAAAGTATGCTCCAGAAAACCTAGATGACGACATTAAAGAA 1800
Db 1857 TTCAGCCCTCTCTACGCTGGACAAAGTATGCTCCAGAAAACCTAGATGACGACATTAAAGAA 1916
Qy 1801 GTGTCCTCCAGCAGATCCTTTGAGAAAGGAGCCTATATCTGTGCCACCCCTCTGGACAGGACA 1860
Db 1917 GTGTCCTCCAGCAGATCCTTTGAGAAAGGAGCCTATATCTGTGCCACCCCTCTGGACAGGACA 1976
Qy 1861 TGC 1863
Db 1977 TGC 1979
RESULT 3
ID ABV23498 standard; cDNA; 2522 BP.
XX ABV23498;
AC ABV23498;
XX 16-SEP-2002 (first entry)
DT Human prostate expression marker cDNA 23489.
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
KW Homo sapiens.
OS XX
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US05171.
XX 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
DR Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX Claim 1; Page 4299-4300; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX Sequence 2522 BP; 642 A; 634 C; 701 G; 543 T; 2 other;
SQ Query Match 100.0%; Score 1863; DB 23; Length 2522;

Db 897 GATAAATAGGCAATTCGGGCTCCACACTTGTGAAGTCCATTTTGAACACCAAGATA 956
QY 841 CTTGTGGAACACATCTTGGAGGTCGGAGATGGTTTAAAGTGGCCATGAACATCTC 900
Db 957 CTTGTGGAACACATCTTGGAGGTCGGAGATGGTTTAAAGTGGCCATGAACATCTC 1016
QY 901 AACAGCGCGGTTTACAGTGGCAGCGCTGCTGGGCTGCTCAAGAGATTGATGAA 960
Db 1017 AACAGCGCGGTTTACAGTGGCAGCGCTGCTGGGCTGCTCAAGAGATTGATGAA 1076
QY 961 ATGACTGCTGAGTACCGCTGCACAGGAACAGTTTAAAGAGGCTCAGTGAATTGGA 1020
Db 1077 ATGACTGCTGAGTACCGCTGCACAGGAACAGTTTAAAGAGGCTCAGTGAATTGGA 1136
QY 1021 TTGATTTCAGAGAAATTTGACATGATGGCTCAGAAGGCTTACGTATGAGAGATGAC 1080
Db 1137 TTGATTTCAGAGAAATTTGACATGATGGCTCAGAAGGCTTACGTATGAGAGATGAC 1196
QY 1081 TACCTCAGCAGGATGCTGGACCAACCTGCTTCCGACTGCTCCATCGAGGAGCC 1140
Db 1197 TACCTCAGCAGGATGCTGGACCAACCTGCTTCCGACTGCTCCATCGAGGAGCC 1256
QY 1141 ATGCTGAAGTGTTCAGCTCCGAGCGCTGCGAGTGTGTGAGTGAGCGCTGCAGATC 1200
Db 1257 ATGCTGAAGTGTTCAGCTCCGAGCGCTGCGAGTGTGTGAGTGAGCGCTGCAGATC 1316
QY 1201 CTCGGGGCTTGGGCTACACAGGAGTATCCGTACAGCGCATCTCGTGACACCGC 1260
Db 1317 CTCGGGGCTTGGGCTACACAGGAGTATCCGTACAGCGCATCTCGTGACACCGC 1376
QY 1261 ATCTCTCTCATCTTCGAGGGAACCAATGAGATTCCTCGGATGTACATCGCCCTGACGGT 1320
Db 1377 ATCTCTCTCATCTTCGAGGGAACCAATGAGATTCCTCGGATGTACATCGCCCTGACGGT 1436
QY 1321 CTCGAGCATCGGCGCGCATCTGACTACAGGATCCATGAGCTTAAACAGGCAAGTG 1380
Db 1437 CTCGAGCATCGGCGCGCATCTGACTACAGGATCCATGAGCTTAAACAGGCAAGTG 1496
QY 1381 AGCAGATCATGATACCTGTTGGCGGAGGCTTCGGGACTCCCTGGGCGCAACTGTGGAC 1440
Db 1497 AGCAGATCATGATACCTGTTGGCGGAGGCTTCGGGACTCCCTGGGCGCAACTGTGGAC 1556
QY 1441 CTGGGGCTACAGGCAACCATGGAGTTGTGACCCGAGTCTTGGGACAGTGCACCAAG 1500
Db 1557 CTGGGGCTACAGGCAACCATGGAGTTGTGACCCGAGTCTTGGGACAGTGCACCAAG 1616
QY 1501 TTTGAGGAGAACACCTACTGCTTGGCGGAGCGCTGAGACACTGCTCGCTTGGC 1560
Db 1617 TTTGAGGAGAACACCTACTGCTTGGCGGAGCGCTGAGACACTGCTCGCTTGGC 1676
QY 1561 AAGACCATCATGAGGAGCAGCTGTTGAGCGGCTGAGACACTGCTCGCTTGGC 1620
Db 1677 AAGACCATCATGAGGAGCAGCTGTTGAGCGGCTGAGACACTGCTCGCTTGGC 1736
QY 1621 TATGAGCATACGGCGCTGCTGTCGGGCGCAGCGCTCATCGCATTTGGGCTCGCAAC 1680
Db 1737 TATGAGCATACGGCGCTGCTGTCGGGCGCAGCGCTCATCGCATTTGGGCTCGCAAC 1796
QY 1681 CAGGACGAGGTTCTCTTGGCCCAACACCTTCTCGGTGGAAGCTTACTTGGAGATCTC 1740
Db 1797 CAGGACGAGGTTCTCTTGGCCCAACACCTTCTCGGTGGAAGCTTACTTGGAGATCTC 1856
QY 1741 TTCAGGCTCTCAGCTGACAGATGATGCTCCAGAAACCTAGATGAGCAGATTAAAGAA 1800
Db 1857 TTCAGGCTCTCAGCTGACAGATGATGCTCCAGAAACCTAGATGAGCAGATTAAAGAA 1916
QY 1801 GTGTCCAGCAGATCTTGGAGAGCGAGCTTATCTGTGCCACCTCTGACAGGACA 1860
Db 1917 GTGTCCAGCAGATCTTGGAGAGCGAGCTTATCTGTGCCACCTCTGACAGGACA 1976
QY 1861 TGC 1863
|||

Db 1977 TGC 1979
RESULT 6
AAS40878
ID AAS40878 standard; cDNA; 2492 BP.
XX
AC AAS40878;
DT 17-DEC-2001 (first entry)
XX cDNA encoding novel human enzyme polypeptide #94.
DE Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ss.
XX Homo sapiens.
OS WO200155301-A2.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US01239.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.

```
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0241617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.

PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 03-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-465566/50.
XX P-PSDB; AAU23008.
DR
DR Novel polypeptides and polynucleotides useful for diagnosing,
XX preventing, treating neural, immune system, muscular, reproductive,
XX pulmonary, cardiovascular, renal, proliferative disorders and cancerous
XX diseases.
XX
XX Claim 4; SEQ ID No 104; 1180pp; English.
XX
XX The present invention relates to the isolation of novel human enzyme
CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS), autoimmune disorders
CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders
CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
CC blood-related disorders (e.g. haemophilia), reproductive disorders
CC (e.g. infertility) and infectious disorders (e.g. influenza). The
CC polynucleotides of the invention can also be used in gene therapy.
CC AAS40785-AAS41684 represent cDNA sequences encoding for the novel human
CC enzyme polypeptides of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2492 BP; 617 A; 621 C; 693 G; 558 T; 3 other;

Query Match 100.0%; Score 1862.6; DB 22; Length 2492;
Best Local Similarity 99.9%; pred. No. 0;
Matches 1862; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGCGGCTGCGGGCTCTTCCTTGGCAGCACACGGCTGCGGCTGCTGCTGCGGGGTCTG 60
Db 67 ATGAGCGGCTGCGGGCTCTTCCTTGGCAGCACACGGCTGCGGCTGCTGCTGCGGGGTCTG 126
Qy 61 GTGGTCTCTACCGCGAACCGGGCTACTGCGCAGCCAGCCGCGCTGTACGAGCTTCGCGC 120
Db 127 GTGGTCTCTACCGCGAACCGGGCTACTGCGCAGCCAGCCGCGCTGTACGAGCTTCGCGC 186
Qy 121 AAAGAGCTTTTCTAGGCAAAATCAAGAAGAAGAGTTTTCCTCCATTTCCAGAAGTTAGC 180
Db 187 AAAGAGCTTTTCTAGGCAAAATCAAGAAGAAGAGTTTTCCTCCATTTCCAGAAGTTAGC 246
Qy 181 CAAGATGAACCTTAATGAATCAATCAGTTCTTGGGACCCGCTGGAAGAAATTTCTTCACTGAA 240
Db 181 CAAGATGAACCTTAATGAATCAATCAGTTCTTGGGACCCGCTGGAAGAAATTTCTTCACTGAA 240
```

247	CAAGATGAACCTTAATGAATCAATCAGTTCTTGGACCGCGTGAATAATCTTCACTGAA	306
241	GAGGTGGACTCCCGAAAAATTGACAGAGGAGGAAAAATCCAGATGAAACTTTGGAGAAA	300
307	GAGGTGGACTCCCGAAAAATTGACAGAGGAGGAAAAATCCAGATGAAACTTTGGAGAAA	366
301	TTGAAGAGCCTTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCCTGGGCTTC	360
367	TTGAAGAGCCTTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCCTGGGCTTC	426
361	TCCAAACCACTGACTCAAGACTAGGGAGATCATCAGCATGGATGGGTCCATCACTGTG	420
427	TCCAAACCACTGACTCAAGACTAGGGAGATCATCAGCATGGATGGGTCCATCACTGTG	486
421	ACCCTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCTGAGGAG	480
487	ACCCTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCTGAGGAG	546
481	CAGAAGGCCAAATACTTGCTTAACTGGCTCCGGGAGCACATTCGAGCCTTCTGCCTC	540
547	CAGAAGGCCAAATACTTGCTTAACTGGCTCCGGGAGCACATTCGAGCCTTCTGCCTC	606
541	ACGGAGCCAGCCAGTGGAGCGATGCAAGCTCAATCCGGAGCAGAGCCACACTAAGTGAA	600
607	ACGGAGCCAGCCAGTGGAGCGATGCAAGCTCAATCCGGAGCAGAGCCACACTAAGTGAA	666
601	GACAAGAGCACTACATCTCAATGGCTCCCAAGGCTCTGGATTACTAATGAGGAGCTGGCC	660
667	GACAAGAGCACTACATCTCAATGGCTCCCAAGGCTCTGGATTACTAATGAGGAGCTGGCC	726
661	AATATTTTACTGTGTTTCCAAAGACTGAGTCTGTTGATTTCTGATGGATCAGTGAAGAC	720
727	AATATTTTACTGTGTTTCCAAAGACTGAGTCTGTTGATTTCTGATGGATCAGTGAAGAC	786
721	AAATACAGCACTTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGAACCCGAA	780
787	AAATACAGCACTTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGAACCCGAA	846
781	GATAAATTAGGCATTTCGGGCTCCAAACACTTCTGAAGTCCATTTTGAACACCAAGATA	840
847	GATAAATTAGGCATTTCGGGCTCCAAACACTTCTGAAGTCCATTTTGAACACCAAGATA	906
841	CCTGTGAAACATCCTTTGGAGAGTCCGAGATGGGTTTAAAGTGGCCATCAACATCCTC	900
907	CCTGTGAAACATCCTTTGGAGAGTCCGAGATGGGTTTAAAGTGGCCATCAACATCCTC	966
901	AACAGCGCGGTTTACGATGGCAGCGTCTGGCTGGGCTGCTCAAGAGATTGATTGAA	960
967	AACAGCGCGGTTTACGATGGCAGCGTCTGGCTGGGCTGCTCAAGAGATTGATTGAA	1026
961	ATGACTGCTGAGTAGCGCTGCACAGGAAACAGTTTAAAGAGGCTCAGTGAATTGGA	1020
1027	ATGACTGCTGAGTAGCGCTGCACAGGAAACAGTTTAAAGAGGCTCAGTGAATTGGA	1086
1021	TTGATTACAGGAAATTTGCATGATGGCTCAGAAGCTTTACGTATGAGAGATGATGCC	1080
1087	TTGATTACAGGAAATTTGCATGATGGCTCAGAAGCTTTACGTATGAGAGATGATGCC	1146
1081	TACCTCACAGCAGGATGCTGGACCAACCTTTCCGAGTGTCCATCGAGGAGCC	1140
1147	TACCTCACAGCAGGATGCTGGACCAACCTTTCCGAGTGTCCATCGAGGAGCC	1206
1141	ATGGTGAAGGTTTCAGCTCCGAGGCGCTGCGAGTGTGTAGTCAGGCGCTGCAGATC	1200
1207	ATGGTGAAGGTTTCAGCTCCGAGGCGCTGCGAGTGTGTAGTCAGGCGCTGCAGATC	1266
1201	CTCGGGGCTTGGGCTACACAAAGGACTATCCGTACGAGCGCATCTGCGTGACACCCG	1260
1267	CTCGGGGCTTGGGCTACACAAAGGACTATCCGTACGAGCGCATCTGCGTGACACCCG	1326
1261	ATCCTTCCTATCTTCAGGGGAACCAATGAGATTCTCCGGATGTACATCCGCTGACGGT	1320
1386	CTCGGGGCTTGGGCTACACAAAGGACTATCCGTACGAGCGCATCTGCGTGACACCCG	1386
1321	CTGCAGCATGCGCGCGCATCTCTGACTACCAAGGATCCATGAGCTTAACAGGCGCAAGTG	1380
1387	CTGCAGCATGCGCGCGCATCTCTGACTACCAAGGATCCATGAGCTTAACAGGCGCAAGTG	1446
1381	AGCACAGTATGATGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGCGCAACTGTGGAC	1440
1447	AGCACAGTATGATGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGCGCAACTGTGGAC	1506
1441	CTGGGCTGACAGGCAACCATGGAGTTGTGCACCCAGTCTTGGGAGCAGTGCCACAAG	1500
1507	CTGGGCTGACAGGCAACCATGGAGTTGTGCACCCAGTCTTGGGAGCAGTGCCACAAG	1566
1501	TTTGAGGAGAACACCTACTGCTTCGGCGGAGCCGTGGAGACACTGCTCTCCGCTTTGGC	1560
1567	TTTGAGGAGAACACCTACTGCTTCGGCGGAGCCGTGGAGACACTGCTCTCCGCTTTGGC	1626
1561	AAGACCATCATGGAGGAGCAGTGTACTGAAGCGGTTGGCAACATCCTCATCAACCTG	1620
1627	AAGACCATCATGGAGGAGCAGTGTACTGAAGCGGTTGGCAACATCCTCATCAACCTG	1686
1621	TATGGCATGACGCGCGTGTGTCGGGGCCAGCGCTCCATCCGCAATTTGGGCTCCGCAAC	1680
1687	TATGGCATGACGCGCGTGTGTCGGGGCCAGCGCTCCATCCGCAATTTGGGCTCCGCAAC	1746
1681	CACGACCGAGGTTCTCTTTGGCCAAACACTTCTCGTGGAGGCTTACTTTCAGAAATCTC	1740
1747	CACGACCGAGGTTCTCTTTGGCCAAACACTTCTCGTGGAGGCTTACTTTCAGAAATCTC	1806
1741	TTTTCAGCTCTCTCAGCTGGACAAAGTATGCTCCAGAAAACCTAGATGAGCAGATTAGAAA	1800
1807	TTTTCAGCTCTCTCAGCTGGACAAAGTATGCTCCAGAAAACCTAGATGAGCAGATTAGAAA	1866
1801	GTGTCCTCAGCATCTTTCAGAGGAGCCTATATCTGTGCCACCTCTGTCAGAGGACA	1860
1867	GTGTCCTCAGCATCTTTCAGAGGAGCCTATATCTGTGCCACCTCTGTCAGAGGACA	1926
1861	TGC 1863	
1927	TGC 1929	
RESULT 7		
AAH24246		
ID	AAH24246 standard; cdna; 2440 BP.	
XX		
AC	AAH24246;	
XX		
DT	11-SEP-2001 (first entry)	
XX	Human oxidoreductase protein ORP-24 encoding cdna.	
DE	Human oxidoreductase protein; ORP; cell proliferative disorder;	
KW	arteriosclerosis; cirrhosis; psoriasis; cancer; endocrine disorder;	
KW	diabetes mellitus; diabetes insipidus; dwarfism; hirsutism; amenorrhea;	
KW	osteoporosis; metabolic disorder; obesity; phenylketonuria;	
KW	hypercholesterolaemia; reproductive cycle defect; infertility;	
KW	ovulatory defect; menstrual cycle defect; endometriosis;	
KW	polycystic ovary disease; spermatogenesis disruption; impotence;	
KW	neurological disorder; epilepsy; stroke; Alzheimer's disease;	
KW	Huntington's disease; Parkinson's disease; Creutzfeldt-Jakob disease;	
KW	meningitis; cerebral palsy; muscular dystrophy; mood disorder; anxiety;	
KW	schizophrenic disorder; infection; autoimmune disorder;	
KW	inflammatory disorder; acquired immunodeficiency syndrome; AIDS; asthma;	
KW	allergy; Crohn's disease; atopic dermatitis; gout; multiple sclerosis;	
KW	rheumatoid arthritis; ulcerative colitis; drug screening;	
KW	toxicity screening; transgenic animal; SNP detection; gene therapy; ss.	
OS	Homo sapiens.	
XX		
Key	Location/Qualifiers	
FT	60..1925	
FT	CDS	
FT	/*tag= a	

FT

/product= "ORP-24"

/function= "Oxidoreductase"

WO200144448-A2.

21-JUN-2001.

07-DEC-2000; 2000WO-US33158.

16-DEC-1999; 99US-0172367.

(INCY-) INCYTE GENOMICS INC.

Yue H, Lal P, Tang YT, Hillman JL, Baughn MR, Azimzai Y, Lu DAM;

WPI: 2001-390245/41.

P-PSDB; AAB73691.

Novel human oxidoreductase protein (ORP) useful for diagnosing, treating and preventing cell proliferative, neurological, viral, reproductive and autoimmune/inflammatory disorders associated with abnormal expression of ORP

Claim 5; Page 134; 136pp; English.

Sequences AAB73668-AAB73694 represent 27 novel human oxidoreductase proteins, designated ORP-1 to ORP-27 respectively, and sequences AAH24223-AAH24249 represent cDNAs encoding ORP-1 to ORP-27. Human ORP proteins and nucleic acids are useful for diagnosing, treating or preventing cell proliferative disorders (e.g. arteriosclerosis, cirrhosis, psoriasis, cancers); endocrine disorders (e.g. diabetes mellitus, diabetes insipidus, dwarfism, hirsutism, amenorrhea, osteoporosis); metabolic disorders (e.g., obesity, phenylketonuria, hypercholesterolemia); reproductive disorders (e.g., infertility, ovulatory and menstrual cycle defects, endometriosis, polycystic ovary disease, disruption of spermatogenesis, impotence); neurological disorders (e.g., epilepsy, stroke, Alzheimer's disease, Huntington's disease, Parkinson's disease, meningitis, Creutzfeldt-Jakob disease, cerebral palsy, muscular dystrophy, mood, anxiety and schizophrenic disorders); viral, bacterial, fungal and parasitic infections; and autoimmune/inflammatory disorders such as acquired immunodeficiency syndrome (AIDS), allergies, asthma, Crohn's disease, atopic dermatitis, gout, multiple sclerosis, rheumatoid arthritis or ulcerative colitis. Human ORP proteins and nucleotides can be used to identify compounds which modulate their activity or expression. ORP nucleic acid sequences may also be used for assessing the toxicity of a test compound, to detect upstream sequences such as promoters and regulatory elements, and to create knock out or knock in animals or transgenic animals to model human disease. Oligonucleotide primers derived from ORP gene sequences may be used to detect single nucleotide polymorphisms (SNPs) and for mapping the naturally occurring genomic sequences. Antibodies specific for ORP proteins may be used in the diagnosis of disorders associated with aberrant ORP expression, in assays to monitor patients being treated with ORP or modulators thereof, and for assessing toxicity of potential drugs.

Sequence 2440 BP; 611 A; 614 C; 682 G; 533 T; 0 other;

Query Match

Best Local Similarity 99.9%; Score 1861.4; DB 22; Length 2440;

Matches 1862; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAGCGGCTCGCGGCTCTTCCTGGCACCACCGGCTGGGCTCGTGCCTGCGCGGCTG 60
 Db 60 ATGAGCGGCTCGCGGCTCTTCCTGGCACCACCGGCTGGGCTCGTGCCTGCGCGGCTG 119
 Qy 61 GTGGTCTCTACCGGACCGCGGCTACTGCGCACCAGCCCGGCTGTACGAGCTTCGCG 120
 Db 120 GTGGTCTCTACCGGACCGCGGCTACTGCGCACCAGCCCGGCTGTACGAGCTTCGCG 179
 Qy 121 AAAGAGCTTTTCTAGGCAAAATCAAGAAGAAGAGTTTCCCATTTCCAGAAAGTTAGC 180
 Db 180 AAAGAGCTTTTCTAGGCAAAATCAAGAAGAAGAGTTTCCCATTTCCAGAAAGTTAGC 239

Qy 181 CAAGATGAATTAATGAATCAATCAGTTCTTTGGACCCGTTGAAAAATTTCTTCACTGAA 240
 Db 240 CAAGATGAATTAATGAATCAATCAGTTCTTTGGACCCGTTGAAAAATTTCTTCACTGAA 299
 Qy 241 GAGGTGGACTCCCGAAAAATTCACAGAGGAAAAATCCAGATGAACATTTGGAGAAA 300
 Db 300 GAGGTGGACTCCCGAAAAATTCACAGAGGAAAAATCCAGATGAACATTTGGAGAAA 359
 Qy 301 TTGAAGAGCGTAGGGCTTTTGGGCTGCAAGTCCACAGAAAGATATGGTGGCTGGCTTC 360
 Db 360 TTGAAGAGCGTAGGGCTTTTGGGCTGCAAGTCCACAGAAAGATATGGTGGCTGGCTTC 419
 Qy 361 TCCACACCATGTACTCAAGACTAGGGAGAGATCATCAGCATGGATGGGTCCATCACTGTG 420
 Db 420 TCCACACCATGTACTCAAGACTAGGGAGAGATCATCAGCATGGATGGGTCCATCACTGTG 479
 Qy 421 ACCCTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTTGGCTGGCACTGAGAG 480
 Db 480 ACCCTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTTGGCTGGCACTGAGAG 539
 Qy 481 CAGAAAGCCAAATATCTTGCCTAAACTGGCGTCCGGGGAGCACATTGCAGCGCTTCTGCTC 540
 Db 540 CAGAAAGCCAAATATCTTGCCTAAACTGGCGTCCGGGGAGCACATTGCAGCGCTTCTGCTC 599
 Qy 541 ACGGAGCCAGCGCTGGGCGGATGCAGCCTCAATCCGGAGCAGAGCCACACTAAAGTAA 600
 Db 600 ACGGAGCCAGCGCTGGGCGGATGCAGCCTCAATCCGGAGCAGAGCCACACTAAAGTAA 659
 Qy 601 GACAAGAGCACTACATCTCTCAATGGCTCCAAAGTCTGGATTAATAAGAGGACTGGCC 660
 Db 660 GACAAGAGCACTACATCTCTCAATGGCTCCAAAGTCTGGATTAATAAGAGGACTGGCC 719
 Qy 661 AATATTTTACTGTGTTGTTGCAAGACTGAGGTCTGTGATTTCTGATGGATCAGTAAAGAC 720
 Db 720 AATATTTTACTGTGTTGTTGCAAGACTGAGGTCTGTGATTTCTGATGGATCAGTAAAGAC 779
 Qy 721 AAAATCACACATTCATAGTAGAAAGACTTTGGTGGAGTCACTAATGGAAACCCGAA 780
 Db 780 AAAATCACACATTCATAGTAGAAAGACTTTGGTGGAGTCACTAATGGAAACCCGAA 839
 Qy 781 GATAAATAGGATTCGGGCTCCAACTTTGTAAGTCCATTTTGAACACCAAGATA 840
 Db 840 GATAAATAGGATTCGGGCTCCAACTTTGTAAGTCCATTTTGAACACCAAGATA 899
 Qy 841 CCTGTGAAACATCTTGGAGAGTCCGAGATGGGTTTAAAGTGCCCATGAACATCCTC 900
 Db 900 CCTGTGAAACATCTTGGAGAGTCCGAGATGGGTTTAAAGTGCCCATGAACATCCTC 959
 Qy 901 AACAGCGCGGTTTCAAGCATGGGACGCTCGTGGCTGGCTCAAGAGATTGATTGAA 960
 Db 960 AACAGCGCGGTTTCAAGCATGGGACGCTCGTGGCTGGCTCAAGAGATTGATTGAA 1019
 Qy 961 ATGACTGCTGAGTACGCTTGCACAGGAACAGTTTAAACAGAGGCTCAGTGAATTTGA 1020
 Db 1020 ATGACTGCTGAGTACGCTTGCACAGGAACAGTTTAAACAGAGGCTCAGTGAATTTGA 1079
 Qy 1021 TTGATTACAGAGAAATTTGCTGCTCAGAGGCTTACGTATGAGAGATGATGACC 1080
 Db 1080 TTGATTACAGAGAAATTTGCTGCTCAGAGGCTTACGTATGAGAGATGATGACC 1139
 Qy 1081 TACCTCACAGCAGGATGCTGACCAACCTTGGCTTTCCGACTGCTCCATCAGGAGCC 1140
 Db 1140 TACCTCACAGCAGGATGCTGACCAACCTTGGCTTTCCGACTGCTCCATCAGGAGCC 1199
 Qy 1141 ATGGTGAAGTCTTCAGCTCCGAGCGGCTGGCAGTGTGTGAGTGGCGCTGCGATC 1200
 Db 1200 ATGGTGAAGTCTTCAGCTCCGAGCGGCTGGCAGTGTGTGAGTGGCGCTGCGATC 1259
 Qy 1201 CTCGGGGCTTTGGGCTACACAAAGGACTATCCGTACGAGCGCATCTCGTGACACCCG 1260
 Db 1260 CTCGGGGCTTTGGGCTACACAAAGGACTATCCGTACGAGCGCATCTCGTGACACCCG 1319

QY	481	CAGAAAGCCAAATAC	TGCGCTAAAC	TGGCGTCCGGGGAGCACAT	TGCGACCTTCTGCCTC	540
Db	527	CAGAAAGCCAAATAC	TGCGCTAAAC	TGGCGTCCGGGGAGCACAT	TGCGACCTTCTGCCTC	586
QY	541	ACGAGCCAGCCAGT	GGGAGCATG	CAGCCTCAATCCGAGCAGAGCCACAC	TAAAGTAA	600
Db	587	ACGAGCCAGCCAGT	GGGAGCCAT	GCAGCCTCAATCCGAGCAGAGCCACAC	TAAAGTAA	646
QY	601	GACAAGACACTAC	ATCCTCAAT	GGCTCCAAAGGTCTG	GAATTACTAAATGGAGACTGGCC	660
Db	647	GACAAGACACTAC	ATCCTCAAT	GGCTCCAAAGGTCTG	GAATTACTAAATGGAGACTGGCC	706
QY	661	AATATTTTACTG	TGTTTCA	AAAGACTGAGGTGCGT	TGATTTCTGATGGATCAGTGAAGAC	720
Db	707	AATATTTTACTG	TGTTTCA	AAAGACTGAGGTGCGT	TGATTTCTGATGGATCAGTGAAGAC	766
QY	721	AAATCACAGCAT	TCATAGT	AGAAAGACCTTTG	TGGTGGAGTCACTAAATGGGAACCCGAA	780
Db	767	AAATCACAGCAT	TCATAGT	AGAAAGACCTTTG	TGGTGGAGTCACTAAATGGGAACCCGAA	826
QY	781	GATAAATTAGGC	ATTCGGGCTC	CAACACTTGTGA	GTCCATTTTGTA	840
Db	827	GATAAATTAGGC	ATTCGGGCTC	CAACACTTGTGA	GTCCATTTTGTA	886
QY	841	CCTGTGGA	AAACATCCTTTG	GAGAGTCCGAGATGG	TTTAAGTGGCCATCAACATCCTC	900
Db	887	CCTGTGGA	AAACATCCTTTG	GAGAGTCCGAGATGG	TTTAAGTGGCCATCAACATCCTC	946
QY	901	AACAGCGCGG	TTTCAGAT	GGCAGTCTG	TGGTCAAGAGATTGATGAA	960
Db	947	AACAGCGCGG	TTTCAGAT	GGCAGTCTG	TGGTCAAGAGATTGATGAA	1006
QY	961	ATGACTGCTG	AGTAGCCCTG	CACAAGGA	AAACAGTTTAAACAGAGCCTCAGTGAATTTGGA	1020
Db	1007	ATGACTGCTG	AGTAGCCCTG	CACAAGGA	AAACAGTTTAAACAGAGCCTCAGTGAATTTGGA	1066
QY	1021	TTGATTCAGG	AGAAATTTGC	ACTGATGGCTC	AGAAAGCTTACGTCATGGAGAGTATGACC	1080
Db	1067	TTGATTCAGG	AGAAATTTGC	ACTGATGGCTC	AGAAAGCTTACGTCATGGAGAGTATGACC	1126
QY	1081	TACCTCACAG	CAGGAGTCTG	CGAACCAACCTGG	CTTCCGACTGCTCCATCGAGGAGCC	1140
Db	1127	TACCTCACAG	CAGGAGTCTG	CGAACCAACCTGG	CTTCCGACTGCTCCATCGAGGAGCC	1186
QY	1141	ATGCTGAAG	GTTCAGT	CCGAGGCGCCTG	GCAGTGTGTGAGTGAAGCGCTGCAGATC	1200
Db	1187	ATGCTGAAG	GTTCAGT	CCGAGGCGCCTG	GCAGTGTGTGAGTGAAGCGCTGCAGATC	1246
QY	1201	CTCGGGGCTT	GGGCTAC	ACAAGGACTAT	CCGTACGAGCGCATCTGCGTGCACACCCG	1260
Db	1247	CTCGGGGCTT	GGGCTAC	ACAAGGACTAT	CCGTACGAGCGCATCTGCGTGCACACCCG	1306
QY	1261	ATCCTCCT	CATCTTCG	AGGGAACCAATG	AGATTCTCCGGATGTACATCGCCCTGACGGGT	1320
Db	1307	ATCCTCCT	CATCTTCG	AGGGAACCAATG	AGATTCTCCGGATGTACATCGCCCTGACGGGT	1366
QY	1321	CTGCAGCAT	CGCGGCGCAT	CTCTGACTAC	CAGATCCATGAGCTTAAACAGGCCAAAGTG	1380
Db	1367	CTGCAGCAT	CGCGGCGCAT	CTCTGACTAC	CAGATCCATGAGCTTAAACAGGCCAAAGTG	1426
QY	1381	AGCACATCAT	TGATACCG	TTGGCCGAGGCTT	CGGAGCTCCCTGGGCGCAACTGTGGAC	1440
Db	1427	AGCACATCAT	TGATACCG	TTGGCCGAGGCTT	CGGAGCTCCCTGGGCGCAACTGTGGAC	1486
QY	1441	CTGGGCTG	ACAGCAACCAT	TGAGTTG	GCACCCACGTCTTGGCGACAGTGCCAAACAAG	1500
Db	1487	CTGGGCTG	ACAGCAACCAT	TGAGTTG	GCACCCACGTCTTGGCGACAGTGCCAAACAAG	1546
QY	1501	TTTGAGGAG	AACACCTACT	GCTTTCGGCCGAGCCGT	GGAGACTGCTGCTCGCTTTGGC	1560
Db	1547	TTTGAGGAG	AACACCTACT	GCTTTCGGCCGAGCCGT	GGAGACTGCTGCTCGCTTTGGC	1606
QY	1561	AAGACCAT	CATGGAGG	CAGCTGTG	TACTAAGCGGGTGGCCAAACATCTCATCAACCTG	1620

[illegible]

RESULT 9

AAC76009

ID AAC76009 standard; cDNA; 2218 BP.

AAC76009;

DT 08-FEB-20

XX
DE
Human OREY OREY 564 20190000

Human: open world, free-form, open-ended

Human; open reading frame; ORF; detection; cytostatic; hepatotropic;
vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
hypotensive; dermatological; immunosuppressive; antinflammatory;
antiviral; antibacterial; antifungal; antirheumatic; antihypoid;
antineuritic; gene therapy; cancer; proliferative disorder; hypertension;
neurodegenerative disorder; osteoarthritis; graft vs host disease;
cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
cholesterol ester storage; systemic lupus erythematosus; infection;
severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
bone damage; cartilage damage; antinflammatory disease; coagulation;
thrombosis; contraceptive; ss.

OS Homo sapiens.

XX
PN
WO200058473-A2XX
PD
05-OCT-2000

XX DE 31

XX

PR 02-APR-1999; 99US-0127636.

PR 30-MAR-2000; 2000US-0540763.

PA (CURA-) CURAGEN CORP.

PI Shinketsu RA. Teach M

XX
DB WPT: 2000-60363/53

DR P-PSDB; AAB41800.
v v

PT Novel nucleic acid

PT neurodegenerative disorders and cardiovascular disease -

XX
XX

PS	Claim 5; Page 2343-2344; 5507pb; English.	
XX	AACT74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397, which represent the human ORFX open reading frames 1 to 3161. The ORFX sequences have activities such as: cytostatic; hepatotropic; vulneryary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic; anticonvulsant; antiarthritic; immunosuppressant; immunostimulant; cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The sequences can be used for determining the presence of or predisposition to, or preventing or treating pathological conditions associated with an ORFX-associated disorder. The nucleic acids can be used to express ORFX proteins in gene therapy vectors. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergies, aplastic anemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to enhance coagulation; to inhibit thrombosis; and as a contraceptive.	
XX	SQ	Sequence 2218 BP; 573 A; 548 C; 609 G; 487 T; 1 other;
	Query Match	90.8%; Score 1690.8; DB 21; Length 2218;
	Best Local Similarity	99.9%; Pred. No. 0;
	Matches 1692; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	170	CAGAAGTTAGCCAAAGATGAACCTTAATGAATCAATCAGTCTTGGGACCCGTGAAAAAT 229
DB	2	CAGAAGTTAGCCAAAGATGAACCTTAATGAATCAATCAGTCTTGGGACCCGTGAAAAAT 61
QY	230	TCCTACTGAAGAGTGGACTCCGAAAAATTGACACAGAGGAAAAATCCAGATGAAA 289
DB	62	TCCTACTGAAGAGTGGACTCCGAAAAATTGACACAGAGGAAAAATCCAGATGAAA 121
QY	290	CTTTGGAGAAATTGAAGAGCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTG 349
DB	122	CTTTGGAGAAATTGAAGAGCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTG 181
QY	350	GCCTGGGCTTCTCCAAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGGT 409
DB	182	GCCTGGGCTTCTCCAAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGGT 241
QY	410	CCATCACTGTGACCTGGAGCGCACCGAGCTATTGGCTCAAGGGATCATCTTGGCTG 469
DB	242	CCATCACTGTGACCTGGAGCGCACCGAGCTATTGGCTCAAGGGATCATCTTGGCTG 301
QY	470	GCATGAGGAGCAGAAAGCCAAATACCTTGCTTAACTGGGCTCGGGGAGCACATTCGAC 529
DB	302	GCATGAGGAGCAGAAAGCCAAATACCTTGCTTAACTGGGCTCGGGGAGCACATTCGAC 361
QY	530	CTTCTGCTCAGCGAGCCAGCTAGTGGGAGCGATGAGCTCAATCCAGGAGATCATCTTGGCTG 589
DB	362	CTTCTGCTCAGCGAGCCAGCTAGTGGGAGCGATGAGCTCAATCCAGGAGATCATCTTGGCTG 421
QY	590	CACCTAAGTGAAGCAAGAAGCAGCTACATCTCAATGGCTCCAAGGCTCGGATTAATG 649
DB	422	CACCTAAGTGAAGCAAGAAGCAGCTACATCTCAATGGCTCCAAGGCTCGGATTAATG 481
QY	650	GAGGACTGGCCAAATATTTTACTGTGTGTTGCAAGAGCTAGGTCGTTGATCTCATGAT 709
DB	482	GAGGACTGGCCAAATATTTTACTGTGTGTTGCAAGAGCTAGGTCGTTGATCTCATGAT 541
QY	710	CAGTGAAGACAAATACACGATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATG 769
DB	542	CAGTGAAGACAAATACACGATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATG 601
QY	770	GGAAACCCGAAAGATAAATTAGGCATTCGGGCTCCAAACACTGTGGAAGTCCATTTGAAA 829
DB	602	GGAAACCCGAAAGATAAATTAGGCATTCGGGCTCCAAACACTGTGGAAGTCCATTTGAAA 661
QY	830	ACACCAAGATACCTGTGGAAAAACATCCTTGGAGAGGTCGGAGATGGGTTTAAAGTGGCCA 889
DB	662	ACACCAAGATACCTGTGGAAAAACATCCTTGGAGAGGTCGGAGATGGGTTTAAAGTGGCCA 721
QY	890	TGAACATCCTCAACAGCGCGGTTTCACATGGGAGCGTGTGGCTGGGCTGGCTCAAGA 949
DB	722	TGAACATCCTCAACAGCGCGGTTTCACATGGGAGCGTGTGGCTGGGCTGGCTCAAGA 781
QY	950	GATTGATTGAATGACTCTGAGTACGCTGCACAAAGAAACAGTTTAAACAGAGGCTCA 1009
DB	782	GATTGATTGAATGACTCTGAGTACGCTGCACAAAGAAACAGTTTAAACAGAGGCTCA 841
QY	1010	GTGAATTTGGATTGATTGAGGAGAAATTTGCACTGATGGCTCAGAGGCTTACGTCATGG 1069
DB	842	GTGAATTTGGATTGATTGAGGAGAAATTTGCACTGATGGCTCAGAGGCTTACGTCATGG 901
QY	1070	AGAGTATGACCTTACCTCACAGCAGGATGCTGGACCAACCTGCTTCCGACTGCTCA 1129
DB	902	AGAGTATGACCTTACCTCACAGCAGGATGCTGGACCAACCTGCTTCCGACTGCTCA 961
QY	1130	TCGAGGAGCCCATGTTGAAGTGTTCAGCTCCGAGGCCCTGGCAGTGTGTGAGTGG 1189
DB	962	TCGAGGAGCCCATGTTGAAGTGTTCAGCTCCGAGGCCCTGGCAGTGTGTGAGTGG 1021
QY	1190	CGCTGAGATCCTCGGGGCTTGGGCTACACAGGAGCTATCCGTACGAGCGCATCTGC 1249
DB	1022	CGCTGAGATCCTCGGGGCTTGGGCTACACAGGAGCTATCCGTACGAGCGCATCTGC 1081
QY	1250	GTGACACCCGATCCTCTCTCATCTTCGAGGGAACCAATGAGATTTCTCCGGATGTACATCG 1309
DB	1082	GTGACACCCGATCCTCTCTCATCTTCGAGGGAACCAATGAGATTTCTCCGGATGTACATCG 1141
QY	1310	CCCTGACGGGCTTGACAGCATGCGGCCGCATCCTGACTACAGGATTCATGAGCTTAAAC 1369
DB	1142	CCCTGACGGGCTTGACAGCATGCGGCCGCATCCTGACTACAGGATTCATGAGCTTAAAC 1201
QY	1370	AGGCCAAAGTGAACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1429
DB	1202	AGGCCAAAGTGAACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1261
QY	1430	GAACTGTGACCTGGGGCTTGACAGGCAACCATGAGTTGTGACACCCAGCTTTCGGGACA 1489
DB	1262	GAACTGTGACCTGGGGCTTGACAGGCAACCATGAGTTGTGACACCCAGCTTTCGGGACA 1321
QY	1490	GTGCCAAAGTGTGAGGAGAACACCTACTGCTTCGGCCGAGCCCTGGAGACTGCTGC 1549
DB	1322	GTGCCAAAGTGTGAGGAGAACACCTACTGCTTCGGCCGAGCCCTGGAGACTGCTGC 1381
QY	1550	TCCGCTTTGGCAAGACCATCATGAGGAGCAGCTGTTGCTTGAAGCGGGTGGCCCAATCC 1609
DB	1382	TCCGCTTTGGCAAGACCATCATGAGGAGCAGCTGTTGCTTGAAGCGGGTGGCCCAATCC 1441
QY	1610	TCATCAACCTGTATGGCATGACGGCGTGTGTCGGGGGCGAGCCGCTCCATCCGATTTG 1669
DB	1442	TCATCAACCTGTATGGCATGACGGCGTGTGTCGGGGGCGAGCCGCTCCATCCGATTTG 1501
QY	1670	GGCTCCGCAACAGCAGCCAGGCTTCTTTCGCAACACTTCTTTCGCTGGAAGCTTACT 1729
DB	1502	GGCTCCGCAACAGCAGCCAGGCTTCTTTCGCAACACTTCTTTCGCTGGAAGCTTACT 1561
QY	1730	TGCAGATCTCTTTCAGCTCTCTCAGCTGGACAAAGTATGTCGCAAGAAACCTAGATGAGC 1789
DB	1562	TGCAGATCTCTTTCAGCTCTCTCAGCTGGACAAAGTATGTCGCAAGAAACCTAGATGAGC 1621
QY	1790	AGATTAGAAGTGTCCACAGCATCTTTCAGAGCGAGCTTATCTGTGCCACCCCTC 1849
DB	1622	AGATTAGAAGTGTCCACAGCATCTTTCAGAGCGAGCTTATCTGTGCCACCCCTC 1681
QY	1850	TGGACGAGCATGC 1863
DB	1682	TGGACGAGCATGC 1695

RESULT 10
AAH15804
ID AAH15804 standard; cDNA; 2400 BP.
XX
AC AAH15804;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:14270.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 95JP-0248036.
PR 27-AUG-1999; 95JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs.
XX
PS Claim 8; SEQ ID 14270; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95993 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 2400 BP; 570 A; 629 C; 685 G; 516 T; 0 other;

Query Match 89.7%; Score 1671.4; DB 22; Length 2400;
Best Local Similarity 92.9%; Pred. No. 0;
Matches 1807; Conservative 0; Mismatches 56; Indels 82; Gaps 2;

QY 1 ATGAGCGGCTGCGGGCTCTTCCTCGCACACCGCTCGGCTCGTCTCGCGGGTCTG 60
|||||
DB 47 ATGAGCGGCTGCGGGCTCTTCCTCGCACACCGCTCGGCTCGTCTCGCGGGTCTG 106
|||||

Tue Sep 9 10:14:38 2003

us-09-945-326-3.rng

QY 1059. TTACGTCATGAGAGTATGACCTACCTACAGCAGGATGCTGGACCAACCTGGCTTCC 1118
DB |||||||
QY 1187 TTACGTCATGAGAGTATGACCTACCTACAGCAGGATGCTGGACCAACCTGGCTTCC 1246
DB |||||||
QY 1119 CGACTGCTCCATCGAGGCGCCATCGTGAAGTGTTCAGTCCGAGGCGCTGGCAGTG 1178
DB |||||||
QY 1247 CGACTGCTCCATCGAGGCGCCATCGTGAAGTGTTCAGTCCGAGGCGCTGGCAGTG 1306
DB |||||||
QY 1179 TGTGAGTAGGCGCTGCAGATCCTCGGGGGCTGGCTACACAAGGACTATCCGTAGA 1238
DB |||||||
QY 1307 TGTGAGTAGGCGCTGCAGATCCTCGGGGGCTGGCTACACAAGGACTATCCGTAGA 1366
DB |||||||
QY 1239 GGCATACATGCGTGCAGCAGGCTCCTCTCATCTTCGAGGGAACCAATGAGATCTCG 1298
DB |||||||
QY 1367 GGCATACATGCGTGCAGCAGGCTCCTCTCATCTTCGAGGGAACCAATGAGATCTCG 1426
DB |||||||
QY 1299 GATGTATATCGGCGCTGCAGGCTGCAGCATGCGCGCGCATCTCTGACTACAGATCA 1358
DB |||||||
QY 1427 GATGTATATCGGCGCTGCAGGCTGCAGCATGCGCGCGCATCTCTGACTACAGATCA 1486
DB |||||||
QY 1359 TGAGCTTAACAGGCCAAAGTGAGCACAGTATGATACGTTGGCGGAGGCTTCGGGA 1418
DB |||||||
QY 1487 TGAGCTTAACAGGCCAAAGTGAGCACAGTATGATACGTTGGCGGAGGCTTCGGGA 1546
DB |||||||
QY 1419 CTCCTGGGCGCAACTGTGACCTGGGGCTGACAGGCAACCATGGAGTTGTGACCCAG 1478
DB |||||||
QY 1547 CTCCTGGGCGCAACTGTGACCTGGGGCTGACAGGCAACCATGGAGTTGTGACCCAG 1606
DB |||||||
QY 1479 TCTTGGGAGAGTGCACCAAGTTTGAGGAGACACTACTGCTTCGGCGGACCGTGA 1538
DB |||||||
QY 1607 TCTTGGGAGAGTGCACCAAGTTTGAGGAGACACTACTGCTTCGGCGGACCGTGA 1666
DB |||||||
QY 1539 GACACTGCTCTCGCTTTGGCAAGACCATCATGAGGAGCAGCTGGTACTGAAGCGGT 1598
DB |||||||
QY 1667 GACACTGCTCTCGCTTTGGCAAGACCATCATGAGGAGCAGCTGGTACTGAAGCGGT 1726
DB |||||||
QY 1599 GGCCAAACATCCTATCAACCTGTATGGCATGAGCGCGTGTCTGCGGGCCAGCGCTC 1658
DB |||||||
QY 1727 GGCCAAACATCCTATCAACCTGTATGGCATGAGCGCGTGTCTGCGGGCCAGCGCTC 1786
DB |||||||
QY 1659 CATCCGATTTGGCTCGGCAACACGACGAGGTTCTCTGGCCAAACACCTTCGCGT 1718
DB |||||||
QY 1787 CATCCGATTTGGCTCGGCAACACGACGAGGTTCTCTGGCCAAACACCTTCGCGT 1846
DB |||||||
QY 1719 GGAAGCTTACTTGAGAACTCTTTCAGCTCTCTGAGTGGCAAGTATGCTCCAGAAA 1778
DB |||||||
QY 1847 GGAAGCTTACTTGAGAACTCTTTCAGCTCTCTGAGTGGCAAGTATGCTCCAGAAA 1906
DB |||||||
QY 1779 CCTAGATGACGAGATTAGAAAGTGTCCAGCAGATCCTTGAAGGAGGCTATATCTG 1838
DB |||||||
QY 1907 CTTAGATGACGAGATTAGAAAGTGTCCAGCAGATCCTTGAAGGAGGCTATATCTG 1966
DB |||||||
QY 1839 TGCCACCCCTCTGGACAGGACATGC 1863
DB |||||||
QY 1967 TGCCACCCCTCTGGACAGGACATGC 1991
DB |||||||

RESULT 11
AAS40882
ID AAS40882 standard; cdna; 1118 BP.

XX AC AAS40882;

XX DT 17-DEC-2001 (first entry)

DE cdna encoding novel human enzyme polypeptide #98.

XX Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ss.

XX Homo sapiens.
OS WO200155301-A2.
PN 02-AUG-2001.
PD 17-JAN-2001; 2001WO-US01239.
PF 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220964.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.


```
|||||
484 CGAAGTGTGGAGCTGGGCTGACAGACACCATGAGTTGTGACCCAGTCTTGGGAC 543
1489 AGTGCCAAACAGTTTGGAGGAACACCTACTGCTTGGCGCGGACCGTGAGACACTGCTG 1548
544 AGTGCCAAACAGTTTGGAGGAACACCTACTGCTTGGCGCGGACCGTGAGACACTGCTG 603
1549 CTCGCGTTTGGCAAGACCATCATGTGAGGAGCAGCTGGTACTGAAGGGGTGCGCAACATC 1608
604 CTCGCGTTTGGCAAGACCATCATGTGAGGAGCAGCTGGTACTGAAGGGGTGCGCAACATC 663
1609 CTCATCAACCTGTATGGCATGACGGCGGTGCTGTCGCGGCGGACCGCTCCATCCGCAAT 1668
664 CTCATCAACCTGTATGGCATGACGGCGGTGCTGTCGCGGCGGACCGCTCCATCCGCAAT 723
1669 GGGCTCCGCAACACGACGACGAGGTTCTTGGGCAACACCTTCTGCGTGAAGCTTAC 1728
724 GGGCTCCGCAACACGACGACGAGGTTCTTGGGCAACACCTTCTGCGTGAAGCTTAC 783
1729 TTGCAGAACTCTTCAAGCCTCTCTCAGCTGGACAGTATGCTCCAGAAAACCTAGATGAG 1788
784 TTGCAGAACTCTTCAAGCCTCTCTCAGCTGGACAGTATGCTCCAGAAAACCTAGATGAG 843
1789 CAGATTAGAAAGTGTCCAGCAGATCCTTTGAGAACGAGCCTATATCTGTGCCACCT 1848
844 CAGATTAGAAAGTGTCCAGCAGATCCTTTGAGAACGAGCCTATATCTGTGCCACCT 903
1849 CTGGACAGGACATGC 1863
904 CTGGACAGGACATGC 918

RESULT 12
AAH08631
ID AH08631 standard; cDNA; 771 BP.
XX AC AAH08631;
XX DT 26-JUN-2001 (first entry)
XX DE Human cDNA clone (5'-primer) SEQ ID NO:5466.
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX OS Homo sapiens.
XX PN EPI074617-A2.
XX PD 07-FEB-2001.
XX PF 28-JUL-2000; 2000EP-0116126.
XX PR 29-JUL-1999; 99JP-0248036.
XX PR 27-AUG-1999; 99JP-0300253.
XX PR 11-JAN-2000; 2000JP-0118776.
XX PR 02-MAY-2000; 2000JP-0183767.
XX PR 09-JUN-2000; 2000JP-0241899.
XX PA (HELI-) HELIX RES INST.
XX PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX Claim 1; SEQ ID 5466; 2537pp + CD ROM; English.
XX The present invention describes primer sets for synthesizing 5602
CC
```

```
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH01366 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 771 BP; 204 A; 187 C; 219 G; 158 T; 3 other;
SQ
Query Match 37.8%; Score 704.6; DB 22; Length 771;
Best Local Similarity 98.9%; Pred. No. 7.7e-179;
Matches 718; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 1 ATGAGCGGCTGGGGCTCTTCTCTGCGCACACCGGCTGCGGCTGCGGCTGCGGGTCTG 60
DB 47 ATGAGCGGCTGGGGCTCTTCTCTGCGCACACCGGCTGCGGCTGCGGCTGCGGGTCTG 106
QY 61 GTGGTCTCTACCGCGAACCGCGGCTACTGCGCACACCGGCTGCGGCTGCGGCTGCGGC 120
DB 107 GTGGTCTCTACCGCGAACCGCGGCTACTGCGCACACCGGCTGCGGCTGCGGCTGCGGC 166
QY 121 AAAGAGCTTTTCTTAGGCAAAATCAAGAAGAAAGAGTTTCCCATTTCCAGAAGTTAG 180
DB 167 AAAGAGCTTTTCTTAGGCAAAATCAAGAAGAAAGAGTTTCCCATTTCCAGAAGTTAG 226
QY 181 CAAGATGAACCTAATGAATCAATCAGTTCTTGGGACCCGCTGAGAAATTTCTCACTGAA 240
DB 227 CAAGATGAACCTAATGAATCAATCAGTTCTTGGGACCCGCTGAGAAATTTCTCACTGAA 286
QY 241 GAGGTGACTCCCGAAAAATTTGACAGAGGAGGAAAAATCCAGATGAAACTTTGGAGAA 300
DB 287 GAGGTGACTCCCGAAAAATTTGACAGAGGAGGAAAAATCCAGATGAAACTTTGGAGAA 346
QY 301 TTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCCAGAGATATGTTGGCCTGGGCTTC 360
DB 347 TTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCCAGAGATATGTTGGCCTGGGCTTC 406
QY 361 TCCAAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGTGGTCCATCACTGTG 420
DB 407 TCCAAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGTGGTCCATCACTGTG 466
QY 421 ACCCTGGCAGCGCACCGAGGCTATTGGCCTCAAGGGGATCATCTTGGCTGGCATGAGGAG 480
DB 467 ACCCTGGCAGCGCACCGAGGCTATTGGCCTCAAGGGGATCATCTTGGCTGGCATGAGGAG 526
QY 481 CAGAAAGCCAAATACTTGGCTTAACATGGCGTCCGGGAGCAGATGCGAGCTTCTGCGCTC 540
DB 527 CAGAAAGCCAAATACTTGGCTTAACATGGCGTCCGGGAGCAGATGCGAGCTTCTGCGCTC 586
QY 541 ACGGAGCCAGCCAGTGGGAGGATGTCAGCTCAATCCGGAGCAGAGCCACACTAAAGTGAA 600
DB 587 ACGGAGCCAGCCAGTGGGAGGATGTCAGCTCAATCCGGAGCAGAGCCACACTAAAGTGAA 646
QY 601 GACAAGAGCAGCTACATCTCAATGGCTCCAAAGTCTGGATTTACTAATGAGGAGCTGGCC 660
DB 647 GACAAGAGCAGCTACATCTCAATGGCTCCAAAGTCTGGATTTACTAATGAGGAGCTGGNC 706
QY 661 AATATTTTACTGTGTTTGCAAGACTGAGGTCTGTTGATCTGTGATGATGATGAGAGAC 720
```

```
|||||
Db 707 AATATTTTACTGGGTTGCNAAGACTGANGTCG-TGATTTCTGATGATCAGTGAAGAC 765
Qy 721 AAAATC 726
Db 766 AAAATC 771

RESULT 13
AAS40879
ID AAS40879 standard; cDNA; 651 BP.
AC AAS40879;
XX
DT 17-DEC-2001 (first entry)
XX
DE cDNA encoding novel human enzyme polypeptide #95.
XX
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ss.
OS Homo sapiens.
XX
XX WO200155301-A2.
PN
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01239.
PF
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.

05-SEP-2000; 2000US-0229509.
05-SEP-2000; 2000US-0229513.
06-SEP-2000; 2000US-0230437.
06-SEP-2000; 2000US-0230438.
08-SEP-2000; 2000US-0231242.
08-SEP-2000; 2000US-0231243.
08-SEP-2000; 2000US-0231244.
08-SEP-2000; 2000US-0231413.
08-SEP-2000; 2000US-0231414.
08-SEP-2000; 2000US-0232080.
08-SEP-2000; 2000US-0232081.
12-SEP-2000; 2000US-0231968.
14-SEP-2000; 2000US-0232397.
14-SEP-2000; 2000US-0232398.
14-SEP-2000; 2000US-0232399.
14-SEP-2000; 2000US-0232400.
14-SEP-2000; 2000US-0232401.
14-SEP-2000; 2000US-0233063.
14-SEP-2000; 2000US-0233064.
14-SEP-2000; 2000US-0233065.
21-SEP-2000; 2000US-0234223.
21-SEP-2000; 2000US-0234274.
25-SEP-2000; 2000US-0234997.
25-SEP-2000; 2000US-0234998.
26-SEP-2000; 2000US-0235484.
27-SEP-2000; 2000US-0235834.
27-SEP-2000; 2000US-0235836.
29-SEP-2000; 2000US-0236327.
29-SEP-2000; 2000US-0236367.
29-SEP-2000; 2000US-0236368.
29-SEP-2000; 2000US-0236369.
29-SEP-2000; 2000US-0236370.
02-OCT-2000; 2000US-0236802.
02-OCT-2000; 2000US-0237037.
02-OCT-2000; 2000US-0237038.
02-OCT-2000; 2000US-0237039.
02-OCT-2000; 2000US-0237040.
13-OCT-2000; 2000US-0239935.
13-OCT-2000; 2000US-0239937.
20-OCT-2000; 2000US-0240960.
20-OCT-2000; 2000US-0241221.
20-OCT-2000; 2000US-0241785.
20-OCT-2000; 2000US-0241786.
20-OCT-2000; 2000US-0241787.
20-OCT-2000; 2000US-0241808.
20-OCT-2000; 2000US-0241809.
20-OCT-2000; 2000US-0241826.
01-NOV-2000; 2000US-0244617.
08-NOV-2000; 2000US-0246474.
08-NOV-2000; 2000US-0246475.
08-NOV-2000; 2000US-0246476.
08-NOV-2000; 2000US-0246477.
08-NOV-2000; 2000US-0246478.
08-NOV-2000; 2000US-0246523.
08-NOV-2000; 2000US-0246524.
08-NOV-2000; 2000US-0246525.
08-NOV-2000; 2000US-0246526.
08-NOV-2000; 2000US-0246527.
08-NOV-2000; 2000US-0246528.
08-NOV-2000; 2000US-0246532.
08-NOV-2000; 2000US-0246609.
08-NOV-2000; 2000US-0246610.
08-NOV-2000; 2000US-0246611.
08-NOV-2000; 2000US-0246613.
17-NOV-2000; 2000US-0249207.
17-NOV-2000; 2000US-0249208.
17-NOV-2000; 2000US-0249209.
17-NOV-2000; 2000US-0249210.
17-NOV-2000; 2000US-0249211.
17-NOV-2000; 2000US-0249212.
17-NOV-2000; 2000US-0249213.
17-NOV-2000; 2000US-0249214.
17-NOV-2000; 2000US-0249215.
```

PR	17-NOV-2000;	2000US-0249216.	
PR	17-NOV-2000;	2000US-0249217.	
PR	17-NOV-2000;	2000US-0249218.	
PR	17-NOV-2000;	2000US-0249244.	
PR	17-NOV-2000;	2000US-0249245.	
PR	17-NOV-2000;	2000US-0249264.	
PR	17-NOV-2000;	2000US-0249265.	
PR	17-NOV-2000;	2000US-0249297.	
PR	17-NOV-2000;	2000US-0249299.	
PR	17-NOV-2000;	2000US-0249300.	
PR	01-DEC-2000;	2000US-0250160.	
PR	01-DEC-2000;	2000US-0250391.	
PR	05-DEC-2000;	2000US-0251030.	
PR	05-DEC-2000;	2000US-0251988.	
PR	05-DEC-2000;	2000US-0256719.	
PR	06-DEC-2000;	2000US-0251479.	
PR	08-DEC-2000;	2000US-0251856.	
PR	08-DEC-2000;	2000US-0251868.	
PR	08-DEC-2000;	2000US-0251869.	
PR	08-DEC-2000;	2000US-0251989.	
PR	08-DEC-2000;	2000US-0251990.	
PR	11-DEC-2000;	2000US-0254097.	
PR	05-JAN-2001;	2001US-0259678.	
XX	(HUMA-)	HUMAN GENOME SCI INC.	
XX	Rosen CA,	Barash SC, Ruben SM;	
XX	WPI;	2001-465566/50.	
DR	P-PSDB;	AAU23009.	
XX	Novel polypeptides and polynucleotides useful for diagnosing,		
PT	preventing, treating neural, immune system, muscular, reproductive,		
PT	pulmonary, cardiovascular, renal, proliferative disorders and cancerous		
PT	diseases		
XX	Claim 4; SEQ ID No 105; 1180pp; English.		
XX	The present invention relates to the isolation of novel human enzyme		
CC	polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences		
CC	encoding them. The enzyme polypeptides of the invention may comprise the		
CC	functional classes of oxidoreductases, transferases, hydrolases, lyases,		
CC	isomerases or ligases. The sequences of the invention are useful in the		
CC	diagnosis, treatment, prevention and/or prognosis of a wide range of		
CC	disorders including hyperproliferative disorders (e.g. cancer),		
CC	immunodeficiency disorders (e.g. AIDS) autoimmune disorders		
CC	(e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),		
CC	metabolic disorders (e.g. phenylketonuria), inflammatory disorders		
CC	(e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),		
CC	blood-related disorders (e.g. haemophilia), reproductive disorders		
CC	(e.g. infertility) and infectious disorders (e.g. Influenza). The		
CC	polynucleotides of the invention can also be used in gene therapy.		
CC	AAU40785-AAU41684 represent cDNA sequences encoding for the novel human		
CC	enzyme polypeptides of the invention.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp.wipo.int/pub/published_pct_sequences.		
XX	Sequence 651 BP; 161 A; 166 C; 194 G; 130 T; 0 other;		
SQ	Query Match	31.4%; Score 585; DB 22; Length 651;	
	Best Local Similarity	100.0%; Pred. No. 9.4e-147;	
	Matches 585; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	ATGAGCGGCTCGGGGCTCTTCTCCGACACACGCGCTCGGCTCGTGCCTCCCGGGTCTG 60	
Db	67	ATGAGCGGCTCGGGGCTCTTCTCCGACACACGCGCTCGGCTCGTGCCTCCCGGGTCTG 126	
Qy	61	GTGCTCTACCGGGAACCGGGGCTACTCGGCACACGCGGCTGTACGAGCTTTCGCC 120	
Db	127	GTGCTCTACCGGGAACCGGGGCTACTCGGCACACGCGGCTGTACGAGCTTTCGCC 186	
Qy	121	AAAGAGCTTTTCTAGCAAAAATCAAGAGAGAAAGTATTTTCCCATTTCCAGAGTTAGC 180	

187	db		187	AAAGAGCTTTCTCCTAGCAAAAATCAAGAAGAAAGAGTATTTCCCATTTCCAGAGTTAGC	246
181	yy		181	CAAGATGAACCTTAATGAATCAATCAGTTCTTGGGACCCGCTGGAAAAATTTCTTCACTGAA	240
247	db		247	CAAGATGAACCTTAATGAATCAATCAGTTCTTGGGACCCGCTGGAAAAATTTCTTCACTGAA	306
241	yy		241	GAGGTGACCTCCCGAAAAAATTTGACCAAGGAGGAAAAATCCAGATGAACCTTTGGAGAAA	300
307	db		307	GAGGTGACCTCCCGAAAAAATTTGACCAAGGAGGAAAAATCCAGATGAACCTTTGGAGAAA	366
301	yy		301	TTGAAGAGCCTTAGGCTTTTGGGCTGCAAGTCCCAAGAAATATGTTGGCTTGGGCTTC	360
367	db		367	TTGAAGAGCCTTAGGCTTTTGGGCTGCAAGTCCCAAGAAATATGTTGGCTTGGGCTTC	426
361	yy		361	TCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCATCACTGTG	420
427	db		427	TCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCATCACTGTG	486
421	yy		421	ACCTGGGAGCGCACCAGGCTATTGGCCCTCAAGGGGATCATCTTGGCTGGCACTGAGGAG	480
487	db		487	ACCTGGGAGCGCACCAGGCTATTGGCCCTCAAGGGGATCATCTTGGCTGGCACTGAGGAG	546
481	yy		481	CAGAAAGCCAAATACTTGCCTAAACTGGCTCCGGGAGACATTTGCAGCCTTCTGCCTC	540
547	db		547	CAGAAAGCCAAATACTTGCCTAAACTGGCTCCGGGAGACATTTGCAGCCTTCTGCCTC	606
541	yy		541	ACGGAGCCAGCCAGTGGGAGCGATCGACGCTCAATCCGGAGCAGA	585
607	db		607	ACGGAGCCAGCCAGTGGGAGCGATCGACGCTCAATCCGGAGCAGA	651
RESULT 14					
AAAX98901					
ID				AAAX98901 standard; cDNA; 772 BP.	
AC				AAAX98901;	
DT				24-SEP-1999 (first entry)	
DE				Human validated cancer cell derived cDNA #223.	
XX				Cancer; human; colon; breast; lung; transmembrane receptor; ATPase;	
XX				integral membrane protein; aspartyl protease; GATA family; wnt family;	
XX				transcription factor; G-protein alpha subunit; protein phosphatase;	
KW				phorbol ester binding protein; diacylglycerol binding protein; trypsin;	
KW				protein kinase; tyrosine phosphatase; developmental signalling protein;	
KW				WW/rsp5/WWP domain; therapy; forensic; genetic mapping; diagnostic;	
KW				detection; treatment; cervical; melanoma; colorectal adenocarcinoma;	
KW				Wilm's tumour; retinoblastoma; sarcoma; myosarcoma; lung carcinoma;	
KW				leukemia; lymphoma; dysplasia; hyperplasia; endometrium; adrenal;	
OS				prostate; ss.	
OS				Homo sapiens.	
XX				WO9933982-A2.	
PN				08-JUL-1999.	
XX					
PD				22-DEC-1998; 98WO-US27610.	
XX					
PF				21-DEC-1998; 98US-0217471.	
XX				23-DEC-1997; 97US-0068755.	
PR				03-APR-1998; 98US-0080664.	
PR				21-OCT-1998; 98US-0105234.	
PR				27-OCT-1998; 98US-0105877.	
XX				(CHIR) CHIRON CORP.	
PA				(HYSE-) HYSEQ INC.	
PA					
XX				Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;	
PI				Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;	
PI				Jones LW, Kassam A, Kennedy GC, Kita D, Labat I;	

0v
1452 AGGCACCATGGAGCTGTGCAACCCAGCTCTGCCCGCACACTTCCACACATTTTCACACACCA 1511

[illegible]

1573

[illegible]

Qy	1570	ATGGAGGACGACTGG--TACTCAAGCGGTGGCCCAACATCTCATCAACCTGTATTGG	1625
Dd	623	ATGGANGCANNINGTGTTNCTNAANGNNNTGGCCAACATNCTCATCAACCTGTATTG	680
 RESULT 15 AAZ15890 ID AAZ15890 standard; cDNA; 835 BP. XX XX AC AAZ15890; XX XX DT 12-OCT-1999 (first entry) XX XX DE Human gene expression product cDNA sequence SEQ ID NO:3359. XX XX Human; gene; gene expression product; diagnosis; therapy; probe; KW detection; mapping; tissue typing; profiling; forensic; cancer; KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.			

PR 31-MAR-1998; 98US-0080114.
PR 03-APR-1998; 98US-0080515.

XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX

PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
PI Escobedo J, García PD, García V, Glese K, Innis MA;

PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
PI Stache-Crain A, Sudduth-Elmore T, Williams M;

XX
DR WPI; 1999-494092/41.

Novel human genes and their expression products which are differentially overexpressed in different cell lines.

PS Claim 1; Page 1608; 2479pp; English.

The present invention describes a library of human polynucleotides

comprising one or more sequences given in SEQ ID NO: 1 to SEQ ID NO: 179. Also described is a method of detecting differentially expressed genes correlated with the cancerous state of a mammalian cell, comprising detecting at least one

CC differentially expressed gene product in a test sample from a cell
CC suspected of being cancerous, where the gene product is encoded by one

of the 32% polynucleotide sequences given in HAZ12332 to HAZ1779. The polynucleotides can be used as a source of primers and probes, which can be used for a variety of purposes, e.g. detection of expression levels

CC mapping, tissue typing or profiling, forensics, genetic analysis and
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides

CC can be used for raising antibodies for experimental, diagnostic and
CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnosis (which may be used to determine function of

CC encoded protein); and to detect differences in expression levels between

CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists.

XX
SQ Sequence 835 BP; 178 A; 212 C; 216 G; 200 T; 29 other;

```
Query Match          26.3%; Score 490.4; DB 20; Length 835;
Best Local Similarity 97.6%; Pred. No. 2.7e-121;
Matches 560; Conservative 0; Mismatches 8; Indels 6; Gaps 6;

QY 1015 TTGGATTGATTC A -GGAGAAATTTGCACTGATGGCTCAGAGGCTTACGTTCATGGAGAG 1073
Db 58 TTGGATTGATTCAGGAGAGAAATTTGCACTGATGGCTCAGAA -GCTTACGTTCATGGAGAG 116

QY 1074 TATGACCTACCTCAGAGGAGGATGCTGACCAACCTGGCTTTCCGACTGCTCCATCGA 1133
Db 117 TATGACCTACCTCAGAGGAGGATGCTGACCAACCTGGCTTTCCGACTGCTCCATCGA 176

QY 1134 GGCAGCCATGGTGAAGTGTTGAGTCCGAGCGCGCTGGCAGTGTGTGAGTGAGCGCGCT 1193
Db 177 GGCAGCCATGGTGAAGTGTTGAGTCCCGA -GCCGCTGGCAGTGTGTGAGTGAGCGCGCT 235

QY 1194 GCAGATCCTCGGGGGCTTGGGCTTACACAAGGGACTATCCGTACGAGCGCATACTGCGTGA 1253
Db 236 GCAGATCCTCGGGGGCTTGGGCTTACACAAGGGACTATCCGTACGAGCGCATACTGCGT-A 294

QY 1254 CACCCGATCCTCCTCATCTTCGAGGAGCAACATGAGATTCCTCGGATGTACATCGCCCT 1313
Db 295 CACCCGATCCTCCTCATCTTCGAGGAGCAACATGAGATTCCTCGGATGTACATCGCCCT 354

QY 1314 GACGGTCTGCAGCATGCCGCGCATCCTGACTACGAGATCCATGAGCTTAAACAGGC 1373
Db 355 GACGGTCTGCAGCATGCCGCGCATCCTGACTACGAGATCCATGAGCTTAAACAGGC 414

QY 1374 CAAAGTGAGCACAGTATGATACCGTTGGCCGGAGGCTTCGGGACTCCCTGGGCGGAAC 1433
Db 415 CAAAGTGAGCACAGTATGATACCGTTGGCCGGAGGCTTCGGGACTCCCTGGGCGGAAC 474

QY 1434 TGTGGACCTGGGGCTGCAGGCAACCATGGAGTTGTGACCCCGAGTCTTGGGACAGTGC 1493
Db 475 TGTGGACCTGGGGCTGCAGGCAACCATGGAGTTGTGACCCCGAGTCTTGGGACAGTGC 534

QY 1494 CAACAAGTTTGGAGAGAACACCTACTGCTTCGCCGGA -CCGTGGAGACAC -TGCCTGCTC 1551
Db 535 CAACAAGTTTGGAGAGAACACCTACTGCTTCGCCGGAACCCGTGGAGACACTTNTTGTTC 594

QY 1552 CGCTTTGGCAAGACCATCATGAGGAGCAGCTGG 1585
Db 595 CGCTTTGGCAAGACCATCATGAGGAGCAACACTTG 628
```

Search completed: September 6, 2003, 17:43:06
Job time : 518.782 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 13:17:56 ; Search time 26 Seconds
(without alignments)
1123.215 Million cell updates/sec

Title: US-09-945-326-2

Perfect score: 3153

Sequence: 1 MSGCGLPLRTAARACRGL.....SOILEKRAYICAHPLDRTC 621

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1360	43.1	656	1	ACDV_MOUSE
2	1356.5	43.0	655	1	ACDV_RAT
3	1345	42.7	655	1	ACDV_BOVIN
4	1308	41.5	655	1	ACDV_HUMAN
5	711	22.5	650	1	Y873_MYCTU
6	691.5	21.9	379	1	ACDB_BACSU
7	652.5	20.7	379	1	ACDA_BACSU
8	630	20.0	389	1	ADCP_MYCTU
9	625.5	19.8	379	1	ACDS_CLOAB
10	609	19.3	389	1	ADCP_MYCLE
11	604.5	19.2	401	1	IVD2_SOLTU
12	604	19.2	432	1	ACDE_HUMAN
13	593.5	18.8	383	1	ACDS_MEGEL
14	587.5	18.6	412	1	IVD1_SOLTU
15	587.5	18.6	423	1	IVD_HUMAN
16	585.5	18.6	424	1	IVD_RAT
17	583	18.5	412	1	ACDS_HUMAN
18	577	18.3	413	1	ACDS_PIG
19	575	18.2	412	1	ACDS_RAT
20	575	18.2	417	1	ACDB_RAT
21	575	18.2	432	1	ACDB_MOUSE
22	572.5	18.2	432	1	ACDB_MOUSE
23	567	18.0	409	1	IVD_ARATH
24	546.5	17.3	409	1	ACDM_HUMAN
25	546.5	17.3	421	1	ACDM_MOUSE
26	542.5	17.2	421	1	ACDM_MOUSE
27	529.5	16.8	421	1	ACDM_PIG
28	525	16.7	419	1	ACDM_DROME
29	523	16.6	421	1	ACDM_RAT
30	507	16.1	413	1	ACDB_MOUSE
31	479	15.2	415	1	ACDB_MOUSE
32	444	14.1	430	1	ACDL_RAT
33	441	14.0	430	1	ACDL_PIG

RESULT 1

ID	ACDV_MOUSE	STANDARD	PRT	656 AA
AC	P50544: Q35289; O55133;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Acyl-CoA dehydrogenase, very-long-chain specific, mitochondrial precursor (EC 1.3.99.-) (VLCAD) (MVLCD).			
GN	ACADVL OR VLCAD.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129/SVJ; TISSUE=Blood;			
RA	Andresen B.S., Lund H., Bross P., Gregersen N.;			
RT	"Mouse very-long-chain acyl-CoA dehydrogenase (VLCAD) gene.";			
RL	Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALB/c; TISSUE=Heart;			
RA	Andresen B., Lund H., Bross P., Corydon M., Gregersen N.;			
RT	"Cloning and characterization of mouse very-long-chain acyl-CoA dehydrogenase.";			
RL	Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE OF 84-656 FROM N.A.			
RC	STRAIN=ICR; TISSUE=Liver;			
RA	MEDLINE=98345418; PubMed=9680378;			
RT	Cox K.B., Johnson K.R., Wood P.A.;			
RL	"Chromosomal locations of the mouse fatty acid oxidation genes Cptla, Cptlb, Cpt2, Acadv1, and metabolically related Crat gene.";			
RN	[4]			
RP	SEQUENCE OF 339-656 FROM N.A.			
RA	Rao G., Krimer D., Krasikov T., Austin C., Skoultschi A.I.;			
RT	Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.			
CC	FUNCTION: ACTIVE TOWARD ESTERS OF LONG-CHAIN AND VERY-LONG CHAIN FATTY ACIDS SUCH AS PALMITOYL-COA AND STEAROYL-COA.			
CC	ACTIVITY: Acyl-CoA + ETF -> 2,3-dehydroacyl-CoA + reduced ETF			
CC	COFACTOR: FAD.			
CC	PATHWAY: Mitochondrial fatty acid beta-oxidation system; first step.			
CC	SUBUNIT: Homodimer (By similarity).			
CC	SUBCELLULAR LOCATION: Mitochondrial inner membrane.			
CC	MISCELLANEOUS: A NUMBER OF STRAIGHT-CHAIN ACYL-COA DEHYDROGENASES OF DIFFERENT SUBSTRATE SPECIFICITIES ARE PRESENT IN MAMMALIAN TISSUES.			
CC	SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 DR EMBL; Y11770; CAA72435.1; -;
 DR EMBL; 271189; CRA94919.2; -;
 DR EMBL; AF017176; AAC31642.1; -;
 DR EMBL; U41497; AAB85185.1; -;
 DR HSSP; Q06319; 1BUC.
 DR SWISS-2DPAGE; P50544; MOUSE.
 DR MGD; MG1:895149; Acadvl.
 DR InterPro; IPR006089; Acyl-CoA_dh.
 DR InterPro; IPR006090; Acyl-CoA_dh_C.
 DR InterPro; IPR006091; Acyl-CoA_dh_M.
 DR InterPro; IPR006092; Acyl-CoA_dh_N.
 DR Pfam; PF00441; Acyl-CoA_dh; 1.
 DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
 DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
 DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
 DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
 DR Oxidoreductase; Flavoprotein; FAD; Fatty acid metabolism;
 KW Mitochondrion; Transient peptide.
 FT TRANSIT 1 41 MITOCHONDRION (BY SIMILARITY).
 FT CHAIN 42 656 ACYL-COA DEHYDROGENASE, VERY-LONG-CHAIN
 SPECIFIC.
 FT DOMAIN 42 483 CATALYTIC.
 FT CONFLICT 339 341 NNG -> GTR (IN REF. 4).
 FT CONFLICT 423 423 C -> W (IN REF. 4).
 FT CONFLICT 427 427 A -> G (IN REF. 4).
 FT CONFLICT 441 441 M -> I (IN REF. 4).
 FT CONFLICT 507 507 G -> A (IN REF. 4).
 FT CONFLICT 532 532 R -> P (IN REF. 4).
 FT CONFLICT 567 567 Q -> K (IN REF. 4).
 FT CONFLICT 570 571 AD -> GG (IN REF. 4).
 FT CONFLICT 573 573 A -> P (IN REF. 4).
 FT CONFLICT 593 593 G -> A (IN REF. 4).
 FT CONFLICT 596 596 T -> A (IN REF. 4).
 FT CONFLICT 612 612 A -> P (IN REF. 4).
 FT CONFLICT 628 628 H -> Q (IN REF. 3).
 FT SEQUENCE 656 AA; 70875 MW; A0110CA5C6CF4F89 CRC64;

Query Match 43.1%; Score 1360; DB 1; Length 656;
 Best Local Similarity 48.2%; Pred. No. 3.7e-82;
 Matches 287; Conservative 106; Mismatches 184; Indels 18; Gaps 9;

QY 35 PVRA----FAKELGKIKKKEVPPEV-SODELNEINQFLGVPVEKFFTEVDSRKIDQ 89
 DB 66 PARAEKSFVQWFKGQLTIDQVFPYPSVLSQAFKELVGPVAFVEVNDPAKND 125
 QY 90 EGPIDETLEKLSGLFLGLOVPEYGGFGFNTWYSLRGEIISM-DGSITVTAAHQAI 148
 DB 126 LEKVEDDTLQGLKELGAFGLQVPSLGGLSNTQYALAEIVGMHDLGVSVTLGAHOSI 185
 QY 149 GLKGIILATEQKAKYPLKASGEHIAFCLTEPASGSDAASIRSRPLSEDKHYILN 208
 DB 186 GFKGILLYGTAKREKYLPRVSGQAALAFCLTEPSSGSDVASIRSSAIPSPCGKIYILN 245
 QY 209 GSKWITNGLANIFTVFAKTEVVD-SDGSVKDKITAFIVDFGVTVNGKPEDKLGIRG 267
 DB 246 GSKIWLSNGLADIFTVFAKTPKDAATGAVREKITAFVVERSGVTVHGLPEKMGKIA 305
 QY 268 SNTCEVHFENKIPVNIIGVGDGPKFVAMNINSGRFSMGVAGLKLRLIEMTAEYAC 327
 DB 306 SNTSEYFVGKVPSENVLGEVGDGPKFVAVNINNGRFGMAATLAGTMKSLIAKAVDHA 365
 QY 328 TRQFNKRLSEGLIOEKALMAQKAYVNESMTYTLFAGMDQPGPDCSIEAAMKVFSS 387
 DB 366 NRTQFGDKIHNFQVIGLEKARMAILQYVTEWMYLSANWDQ-GFKDFQIEAIAIKFCS 424
 QY 388 EAAWQCVSEALQILGLGYTRYDPPYERILDRTRILLIFEGTEILRMVYALTGLQHGRI 447
 DB 425 EAAWKADECIIQMGGMGFMKPGVVERVLRDIRIFRIFEGANDILRLFVALQGMCKGKE 484

QY 448 LTTRIHELKO--AKVSTVMDTVGRRRLRSLGRTVDLGTGNHGHWVHPSLADSANKFEENT 505
 DB 485 LTGLGNALKNPFNGVGLLMGEAGKQLRRRTGIGSGLSLS---GIVHPELSRSGELAVQAL 541
 QY 506 YCFGRVTETLLRFGKTIMEEQVLVKRVANILNLVGMTAVLSRASHSIRILGLRHHDEHV 565
 DB 542 DQFATVVEAKLVKHKHGIIVNEQFLQLRLADGADLYAMVVVLSRASRSLSEGYPTAQHEK 601
 QY 566 LLANTFCVEA---YLQNLFSLSQLDKYAPENLDEQIKKVSQOILEKRAYICAHPL 617
 DB 602 MLCDSWCIEAATRIRENNASLQSSPQH--QELFNFRFSISKAMVENGLVTGNPL 654

RESULT 2

ID ACYLV_RAT STANDARD; PRT; 655 AA.
 AC P45953;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Acyl-CoA dehydrogenase, very-long-chain specific, mitochondrial
 DE precursor (EC 1.3.99.-) (VLCAD).
 GN ACADVL OR VLCAD.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=WiStar; TISSUE=Liver;
 RX MEDLINE=94308174; PubMed=80345667;
 RA Aoyama T., Ueno I., Kamijo T., Hashimoto T.;
 RT "Rat very-long-chain acyl-CoA dehydrogenase, a novel mitochondrial
 RT acyl-CoA dehydrogenase gene product, is a rate-limiting enzyme in
 RT long-chain fatty acid beta-oxidation system. cDNA and deduced amino
 RT acid sequence and distinct specificities of the cDNA-expressed
 RT protein." *Biol. Chem.* 269:19088-19094(1994).
 RL SwissProt: P45953.
 CC -!- FUNCTION: ACTIVE TOWARD ESTERS OF LONG-CHAIN AND VERY-LONG CHAIN
 CC FATTY ACIDS SUCH AS PALMITOYL-COA AND STEAROYL-COA.
 CC -!- CATALYTIC ACTIVITY: Acyl-CoA + ETF -> 2,3-dehydroacyl-CoA + reduced
 CC ETF.
 CC -!- COFACTOR: FAD.
 CC -!- PATHWAY: Mitochondrial fatty acid beta-oxidation system; first
 CC step.
 CC -!- SUBUNIT: Homodimer.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
 CC -!- MISCELLANEOUS: A NUMBER OF STRAIGHT-CHAIN ACYL-COA DEHYDROGENASES
 CC OF DIFFERENT SUBSTRATE SPECIFICITIES ARE PRESENT IN MAMMALIAN
 CC TISSUES.
 CC -!- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; D30647; BAA06331.1; -;
 DR PIR; A54872; A54872.
 DR HSSP; Q06319; 1BUC.
 DR InterPro; IPR006089; Acyl-CoA_dh.
 DR InterPro; IPR006090; Acyl-CoA_dh_C.
 DR InterPro; IPR006091; Acyl-CoA_dh_M.
 DR InterPro; IPR006092; Acyl-CoA_dh_N.
 DR Pfam; PF00441; Acyl-CoA_dh; 1.
 DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
 DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
 DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
 DR PROSITE; PS00073; ACYL_COA_DH_2; 1.

SEQUENCE FROM N.A.
TISSUE=Heart;
Zhang X., Liu W., Zhu H., Sun X.;
Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: ACTIVE TOWARD ESTERS OF LONG-CHAIN AND VERY-LONG CHAIN
FATTY ACIDS SUCH AS PALMITOYL-COA AND STEAROYL-COA.
-!- CATALYTIC ACTIVITY: Acyl-CoA + ETF -> 2,3-dehydroacyl-CoA + reduced
ETF.
-!- COFACTOR: FAD.
-!- PATHWAY: Mitochondrial fatty acid beta-oxidation system; first
step.
-!- SUBUNIT: Homodimer (By similarity).
-!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
-!- MISCELLANEOUS: A NUMBER OF STRAIGHT-CHAIN ACYL-COA DEHYDROGENASES
OF DIFFERENT SUBSTRATE SPECIFICITIES ARE PRESENT IN MAMMALIAN
TISSUES.
-!- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/ or
send an email to license@isb-sib.ch).

EMBL; U30817; AAA74051.1; -
HSSP; O06319; 1BUC.
InterPro: IPR006089; Acyl-CoA_dh.
InterPro: IPR006090; Acyl-CoA_dh_C.
InterPro: IPR006091; Acyl-CoA_dh_M.
InterPro: IPR006092; Acyl-CoA_dh_N.
Pfam; PF00441; Acyl-CoA_dh; 1.
Pfam; PF02770; Acyl-CoA_dh; 1.
Pfam; PF02771; Acyl-CoA_dh; 1.
PROSITE; PS00072; ACYL-COA-DH_1; 1.
PROSITE; PS00073; ACYL-COA-DH_2; 1.
Oxidoreductase; Flavoprotein; FAD; Fatty acid metabolism;
Mitochondrion; Transit peptide.
TRANSIT 1 40 MITOCHONDRION (BY SIMILARITY).
CHAIN 41 655 ACYL-COA DEHYDROGENASE, VERY-LONG-CHAIN
SPECIFIC.
DOMAIN 41 482 CATALYTIC.
SEQUENCE 655 AA; 70520 MW; F9DC06285023CFC0 CRC64;

Query Match 42.7%; Score 1345; DB 1; Length 655;
Best Local Similarity 47.3%; Pred. No. 3.6e-81;
Matches 285; Conservative 110; Mismatches 194; Indels 14; Gaps 9;

QY 23 STANRLRLRTSPVRAFAKELFLGKIKKEVFPFV-SQDELNEINQFLGVPKFFTEE 81
DB 57 SEASTREKANSVSKSFVAGTFFGQLTDDQVFPVPSVLNEDQTQFLKELVGPVTFEEV 116
QY 82 VDSRKIDQEGKIPDETLEKLSGLFLGQVPEYGGGLGFSNTWYSLRGITISM-DGSITV 140
DB 117 NDAAKNDLERVEETTMQGLKELGAFGLQVNPENLGGVGLCTQYARLVEIVGMYDLGVGI 176
QY 141 TLAHQAGLIGLILAGTEBQAKYLPKLASGEHTAAFCITPASPASDAASRSTRATLSE 200
DB 177 VLGAHQSIGFGKILLFCTKAQKEKYLPLKASGETTAAFCITPSPSGSAAASRSPSP 236
QY 201 DKHYILNGSKVITNGGLANITVFATKEVD-SGDSVKDKITAFIVERDGGVTNGKP 259
DB 237 CGKYTLNGSKHISNGGLADITVFATKPTVDTATGAVKKEKITAFVVERSGVTHGPP 296
QY 260 EDKLGIRGNTCEVHFKTIPVENTLGEVGGFKYAMNINLSGRFSMGVYVGLLKRLL 319
DB 297 EKKMGIKASNTAEVYFDGVRVPAENVLGEVGGFKYAMHILNNGRFGMAALAGTMKGI 356
QY 320 EMTAEACTRKQFNKRSLSEGLLOEKFLMAQKAYVMESMTYLTACMLDQGFDPDCSEA 379
DB 357 AKAVDGAARTQEGEKIHNFGLIEKLARMALQVYTESMAYMSANNDD-GSTDQFQIEA 415

OXidoreductase; Flavoprotein; FAD; Fatty acid metabolism;
Mitochondrion; Transit peptide.
TRANSIT 1 40 MITOCHONDRION.
CHAIN 41 655 ACYL-COA DEHYDROGENASE, VERY-LONG-CHAIN
SPECIFIC.
DOMAIN 41 482 CATALYTIC.
SEQUENCE 655 AA; 70749 MW; E808DEB084595D7 CRC64;

Query Match 43.0%; Score 1356.5; DB 1; Length 655;
Best Local Similarity 46.4%; Pred. No. 6.3e-82;
Matches 295; Conservative 108; Mismatches 192; Indels 41; Gaps 12;

QY 9 RTTAAARACGLVSTANRLRL-----TSPVRA-----FAKELFLGKIKKEV 53
DB 32 RPTSAQR-----LYASEATQVLEKPEKTSADASTREKPARAESKFAVGMFKGLTDDQ 87
QY 54 FPPEYSODELNE-----INQFLGVPKFFFEVDSRKIDQEGKIPDETLEKLSGLFG 108
DB 88 FPVPSV-----LNEGQTFKLKELGVPVAFREVDNPAKNDSEKLEKVEEDTLOGLKELGAF 143
QY 109 LQVPEYGGGLGFSNTWYSLRGITISM-DGSITVTLAAHQAGLIGLILAGTEBQAKYLP 167
DB 144 LQVPELGGGLGSLNTQYARLAEIVGMHDLGVSITLGAHQSGTGFKGILLYGTAKQEKYLP 203
QY 168 KLASGEHTAAFCITPASPASDAASRSTRATLSEDKKHVILNGSKVITNGGLANITVF 227
DB 204 RVASGQALAAFCITPSPSGSDVASRSPSPCGKYTLNGSKHISNGGLADITVF 263
QY 228 KTEVD-SGDSVKDKITAFIVERDGGVTNGKPEDKLGIRGNTCEVHFKTIPVBNIL 286
DB 264 KTIKDAATGAVKEKITAFVVERSGVTHGLPEKKMGIKASNTSEVYFDGVKVPANVL 323
QY 287 GEVGDGFKVAMILNGSRFSGSVVAGLLKRLIEMTAFYACTRKQFNKRSLSEGLLOEK 346
DB 324 GEVGDGFKVAVILNNGRFGMAATLGAETPMKAIILAKAVDHAATNRTQFGDKIHNFVQEK 383
QY 347 ALMAQKAYVMESMTYLTACMLDQGFDCSEAAKVVFSSEAAQCVSEALQITGLGV 406
DB 384 ARMAILQVYTESMAYMLSANNDQ-GFKDFQIEAALSIFGSEAAKWTDCIQIMGMGF 442
QY 407 TRDYPVETRLTRILLFEGTNEILLMYALTLQIAGRIITRIHELKQ---AKYSTVM 464
DB 443 MKRQGVRLDIRIFRIFEGTNDILFLVALQCMQDKGELTGLGNALKNPLGNVGLLI 502
QY 465 DTVGRRLRDLGRTVDLGLTGNHGVHPSLADSANKFEENTYCFGRVETLLRFGKTI 524
DB 503 GEASQLRRRTGIGSLSL---GIVHPELSRSELAVQALEQFATVVEAKLHKHKGIV 559
QY 525 EQLVLRKVANILNLYGMTAVLSRASRSTRIGLRNHDHVELLANTFCVEA---YLONLF 581
DB 560 NEQFLQLQALDCAIDLXAMVYVLSRASRSISEGYPTAQHEKMLCDSNCIEAATRIENMA 619
QY 582 SLISQDKYAPENLDEQIKVSOQLIKRAXICAHPL 617
DB 620 SLQSNPQ--QQLFNFRFSISKAMVENGGLVTSNPL 653

RESULT 3
ACDV_BOVIN
ID ACDV_BOVIN STANDARD; PRT; 655 AA.
AC P48618;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acyl-CoA dehydrogenase, very-long-chain specific, mitochondrial
DE precursor (EC 1.3.99.-) (VLCD).
GN ACADVL OR VLCD.
OC Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
RN [1]

323 VQGEAVASKISFTIAATNVADVAVVELSSQMADE-GRNDTRIEAAALAKLWSSSEMAGLVGD 38

396 EAQILQGLG--TRDPYPERILROTRILLIFEGNEILRMVYALTG---L 441

382 ELLOIRGGRGVYETAESLAARGERAVPVQMVDRLRINRIFEGSSSEIMRLIIAREAVDAHL 41

442 QHAGRLITRTHLHQAKQVSTVMDTVGRLRDSLGRIVDLGLTGNHGCVHPSLA----- 495

442 TAAGDLANPKA-DLRQKAAAA-----GASGFYAKWLPKLVFEGGQL 482

496 -----DSANKFEENTYCFGRVETILLRFRGKTIMEEQVLKRVANIL 537

483 PTTYREFGALATHLRFVERSSKRLARNTF-YG-----MARQWASLEKKQFGLGRIVDIG 535

538 INLYGMTAVLSRASRSIRIGLRNHDHVELLANTFC-----VEAYLQNLVSLSLQDKYAP 591

536 AELFAISAAVCVRAEQRTADPVEGQAYELAEAFCCQATLRUEALFDALWS----- 586

592 ENLDEQIKKVSQOILEKR 609

587 -NTDSIDVRLANDVLEGR 603

RESULT 6

ACDB_BACSU STANDARD; PRT: 379 AA.

ID ACDB_BACSU

AC P45857; 1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE ACV1-CoA dehydrogenase (EC 1.3.99.-).

GN MMGC.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

[1]

RN SEQUENCE FROM N.A.

RC STRAIN=168 / MB24;

RC MEDLINE=96326319; PubMed=8759838;

RA Bryan E.M., Beall B.W., Moran C.P. Jr.;

RT "A sigma E dependent operon subject to catabolite repression during

RT sporulation in Bacillus subtilis";

RL J. Bacteriol. 178:4778-4786(1996).

[2]

RN SEQUENCE FROM N.A.

RC STRAIN=168 / JH642;

RC MEDLINE=97124195; PubMed=8969508;

RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,

RA Kobayashi Y.;

RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of

RT the Bacillus subtilis genome containing the skin element and many

RT sporulation genes.";

RL Microbiology 142:3113-3111(1996).

[3]

RN SEQUENCE FROM N.A.

RC STRAIN=168;

RC MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,

RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,

RA Choi S.K., Codani J.J., Conterton I.F., Cummings N.J., Daniel R.A.,

RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,

RA Entian K.D., Errington J., Fabret C., Ferrazi A., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,

RA Ghm S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,

RA Gutseppl G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,

RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,

RA Medina N., Mellado R.P., Mizuno N., Mostl D., Nakai S., Noback M.,

```

RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate F.,
RA Sato T., Scanlan E., Schleich S., Schroeter K., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serrero P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weltzienegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT *The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis*.
RL Nature 390:249-256(1997).
CC
CC -I- CATALYTIC ACTIVITY: Acyl-CoA + ETF = 2,3-dehydroacyl-CoA + reduced
CC ETF.
CC
CC -I- COPACITOR: FAD (BY SIMILARITY).
CC
CC -I- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U29084; AAB09615.1; -
DR EMBL; D84432; BAA12589.1; -
DR EMBL; Z99116; CAB14346.1; -
DR PIR; D69558; D69658.
DR HSP; Q06319; 1BUC.
DR Subtilist; BG11321; mmgC.
DR InterPro; IPR006089; Acyl-CoA_dh.
DR InterPro; IPR006090; Acyl-CoA_dh_C.
DR InterPro; IPR006091; Acyl-CoA_dh_M.
DR InterPro; IPR006092; Acyl-CoA_dh_N.
DR Pfam; PF00441; Acyl-CoA_dh; 1.
DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00073; ACYL_COA_DH_1; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
DR OXIDOREDUCTASE; Flavoprotein; FAD; Complete proteome.
FT CONFLICT 97 97 M > L (IN REF. 2).
FT CONFLICT 171 171 D > A (IN REF. 2).
FT CONFLICT 329 329 MISSING (IN REF. 2).
FT SEQUENCE 379 AA; 41003 MW; 783FF99296DD3F2D CRC64;
Query Match 21.9%; Score 691.5; DB 1; Length 379;
Best Local Similarity 39.5%; Pred. No 2e-38;
Matches 159; Conservative 68; Mismatches 131; Indels 45; Gaps 8;
QY 36 VRAFAKELFGKTKKKEVFPFPEVSDNELNEIQLFGLPVKEFTBEVDSRKIDQEGKIPD 95
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 14 VRDFA-----RKEIAPAAEI-----NEKTDPEPF 37
QY 96 ETLEKLSLGLFGLQPEVGGVGLGFNTMY-SRLGEIISNDSGITVTTLAAHQAIKGI 154
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 38 QLIKMKKIGLGMIPVPGIYGAGADVSYILAIHISRAAVGVILSVHTSVGTNPIM 97
QY 155 LAGTEQKAKYLPKLASGEGHIAAFCLITEPASGSDAASRSRATLSEDKHYILNGSKYWI 214
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 98 YFGNEQKMYIPNLASGDHLGAFALTEPHSGSDAGSLRTTAIKNGK--YLLNGSKIFI 155
QY 215 TNGGLANITVFAKTEVVDSDGSKVDKTAFTIVERDFGVGTNGKPEDKLGIRSGNCEVH 274
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 156 TNGGAADIYITFALY---DPDQG-RHGISAFIVEKNTPGFTVGKKERKGLGVSNTELI 211
QY 275 FENTIPVENILGEGDGFVKVAMNTLNSRRSGMGSWAGLLKRLIEWTAETACTRKQPNK 334
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 212 FDNAEVPENLLCKEGDGFHTAMNLNVRIGIAQAALGAAEALEHADVAVKQRFGR 271

```

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>) or send an email to license@isb-sib.ch.

EMBL; Z49782; CAA89868.1; -
EMBL; Z99123; CAB15745.1; -
PIR; S55421; S55421.
HSSP; Q06319; 1BUC.
DR Subtilist; BG11239; acda.
DR InterPro; IPR006089; Acyl-CoA_dh.
DR InterPro; IPR006090; Acyl-CoA_dh_C.
DR InterPro; IPR006091; Acyl-CoA_dh_M.
DR InterPro; IPR006092; Acyl-CoA_dh_N.
DR Pfam; PF00441; Acyl-CoA_dh; 1.
DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
KW Oxidoreductase; Flavoprotein; FAD; Complete proteome.
SQ SEQUENCE 379 AA; 41446 MW; 4D09861D59718EF9 CRC64;

Query Match 20.7%; Score 652.5; DB 1; Length 379;
Best Local Similarity 42.1%; Pred. No. 7.4e-36;
Matches 155; Conservative 60; Mismatches 140; Indels 13; Gaps 7;

QY 74 VEKFTTEV--DSRKIDQEKIPDETLEKLSGLFGLQVPEEYGGGLFSNTMY-SRLGE 130
DB 16 VRFDAHEVAPTAAERDEQERDFRELFREMANLGTGPWPEDYGGIGSDYLAVIAVEE 75
QY 131 IISMDGSTIVTAAHQATGLKGIILAGTEQAKYLPKLSAGEHIAAFCLTEPAGSDAA 190
DB 76 LSKVCASGTVLSAHLSICSWPLFAFGVEEQKTEVLTQLGKIGAFALTEAGSDAG 135
QY 191 SIRSRAITSEDKKHYLNGSKVWITNGLANITFTVFAKTEVVDSDGSKDK-ITAFIVYER 249
DB 136 SMKTTAERIGD--DYVLNGSKVFTITNGGVADIYVFAVT-----DPEKKKKGVTAFIGEK 188
QY 250 DFGGVNKGPKDLGRSNTCEVHFENTKIPVENILGEVGDGFKVAMNINLSGRFSMGS 309
DB 189 DEGFFTGKKEKLGIRSPTEIMFEDCVVPASRLGEEGEGFKIAMKTLDDGGRNGIAA 248
QY 310 VVAGLLKRLIEMTAETACTKRFNKRLESEGLIQEKFALMAKAVVMESMYLYIFAGLDQ 369
DB 249 QAVGTAQALDAALQAKERKQFGKSIARQOQIAEFLADMATMIEASRLITY-QAAWLES 307
QY 370 PGFPDCSIEAMVKVFSSEAAQCYSALQIILGGLGYTRDYPYERILRDRTRILLIFEGTN 429
DB 308 SGLP-YGKASAMSKLMAGOTAMKVTEAVQIFGGYGYTKDYPVRYMRDAKITQIYEGTQ 366
QY 430 EILRMVIA 437
DB 367 EIQRLVIS 374

RESULT 8
ACDP_MYCTU STANDARD; PRT; 389 AA.
ID ACDP_MYCTU
AC P96879;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable acyl-CoA dehydrogenase fadE25 (EC 1.3.99.-).
GN FADE25 OR ACD OR RV3274C OR MT3374 OR MTCV71.14C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
Bisshai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + ETF -> 2,3-dehydroacyl-CoA + reduced
ETF.
CC -1- COFACTOR: FAD (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ACYL-CoA DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>)
or send an email to license@isb-sib.ch.
CC
EMBL; Z92771; CAB07077.1; -
EMBL; AE007147; AAK47715.1; -
PIR; C70979; C70979.
HSSP; Q06319; 1BUC.
DR TIGR; MT3374; -
DR Tuberculist; RV3274C; -
DR InterPro; IPR006089; Acyl-CoA_dh.
DR InterPro; IPR006090; Acyl-CoA_dh_C.
DR InterPro; IPR006091; Acyl-CoA_dh_M.
DR InterPro; IPR006092; Acyl-CoA_dh_N.
DR Pfam; PF00441; Acyl-CoA_dh; 1.
DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
KW Hypothetical protein; Oxidoreductase; Flavoprotein; FAD;
KW Complete proteome.
SQ SEQUENCE 389 AA; 41723 MW; 2E77D0F28A9C8E08 CRC64;

Query Match 20.0%; Score 630; DB 1; Length 389;
Best Local Similarity 38.0%; Pred. No. 2.3e-34;
Matches 147; Conservative 77; Mismatches 149; Indels 14; Gaps 9;

QY 52 EVFPPEVSQDELNEINQFLGPVEKFFTEEDVDSRKIDQEKIPDETLEKLSGLFGLQV 111
DB 11 DLFLKLP-EHDEMRSAIRAL--AEKEIAPH--AAVDEKARPEEALVALNSSGNVAHI 65
QY 112 PEYGGGLGFSNTMYS-RLGEIISMDGSTIVTAAHQATGLKGIILAGTEQAKYLPKLA 170
DB 66 PEYGGQAGDSVATCIVTEEVARDASASLIPAVNK-LCTMGLILRGSEELKKQVLPALA 124
QY 171 SGEHTAFCLETPPAGSDAASTRATLSEDKKHVILNGSKVWITNGLANITFTVFAKTE 230
DB 125 AEGAMASYALSEREAGSDAASMTTRA--KADGDHMLNGKAWITNGKSWITVYMA--- 179
QY 231 VVDSGSKVDKITAFIVERDFGVTNGKPEDKIGIRGNSNTCEVHFENTKIPVENILGEV 290
DB 180 VTDPRGA-NGISAPMVKHDDGFTVGPKERKLGKSGPTTELYENCRIPIGDIIGEPG 238

```

QY 291 DGFVAMNINLGRSMGVSAGLLKRLIEMTAETACTRKOFNKLSEFGLIOEKPALMA 350
DB 239 TGFKTALATLDHTRPTIGAQAQVGAQALDAAIYTKDRQFGSISTFAQVFMFLADMA 298
QY 351 QKAYVMSWYLLTAGMLDQGFPCDSTEAAVMKVFSEAAQVCSEALQILGLGIGYTRDY 410
DB 299 MKVEARLWYSAARAER-GEPLGFIISAASKCFASDVAMEVTTDAVQLFGGAGYTTDF 357
QY 411 PYERILDRTRILLFEETNELRMVIA 437
DB 358 PVERFMDAKITQIYECTNOLQRVMS 384

RESULT 9
ACDS_CLOAB STANDARD; PRT; 379 AA.
AC P52042;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Acyl-CoA dehydrogenase, short-chain specific (EC 1.3.99.2)
DE (SCAD) (Butyryl-CoA dehydrogenase).
DE BCD OR CAC2711.
GN Clostridium acetobutylicum.
OS Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=96236011; PubMed=8655474;
RA Boynton Z.L., Bennett G.N., Rudolph F.B.;
RT "Cloning, sequencing, and expression of clustered genes encoding
RT beta-hydroxybutyryl-coenzyme A (CoA) dehydrogenase, crotonase, and
RT butyryl-CoA dehydrogenase from Clostridium acetobutylicum ATCC
RT 824.";
RL J. Bacteriol. 178:3015-3024(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
CC -1- CATALYTIC ACTIVITY: Butanoyl-CoA + ETF = 2-butenoyl-CoA + reduced
CC ETF.
CC -1- COFACTOR:- FAD.
CC -1- PATHWAY: BUTYRATE/BUTANOL-PRODUCING PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U17110; AAA95968.1;
CC DR EMBL; AE007768; AAK80657.1;
CC DR PIR; F97233; F97233.
CC DR PIR; T47262; T47262.
CC DR HSP; Q06319; 1BUC.
CC DR InterPro; IPR006089; Acyl-CoA_dh.
CC DR InterPro; IPR006090; Acyl-CoA_dh.C.
CC DR InterPro; IPR006091; Acyl-CoA_dh.M.
CC DR InterPro; IPR006092; Acyl-CoA_dh.N.
CC Pfam; PF00441; Acyl-CoA_dh; 1.

```

```

DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
KW Oxidoreductase; Flavoprotein; FAD; Complete proteome.
SQ SEQUENCE 379 AA; 41386 MW; 4BC50A1991BD9FB2 CRC64;

Query Match 19.8%; Score 625.5; DB 1; Length 379;
Best Local Similarity 39.0%; Pred. No. 4.5e-34;
Matches 144; Conservative 73; Mismatches 137; Indels 15; Gaps 8;

QY 74 VEKFFTEEVD--SRKIDQEGKIPDETLEKLSGLFGLQVPEYGGGLGFSNTMY-SRLGE 130
DB 16 VREFAENEVKPIAAEIDETERFPMENVKMGQYGMGIPFSEYGGAGGVLSYIAVEE 75
QY 131 IISMDGSITVTTLAAHQAIGLKGIILAGTBEQAKYLPKLASGEHTAAFLCTPAGSDAA 190
DB 76 LSKVCGTTGCVILSAHTSLCASLINEHGTTEQKQYLVPLAKGEKICAYGLTEPNAGTDSG 135
QY 191 SIRSRAITLSEDKKHVILNGSKVWITNGGLANITFTVFAKTEVVDSDGSKDKITAFIVERD 250
DB 136 AOOTVAVEGD--HYVINGSKIFITNGGVADTFVIFAMT--DRTKGTGK-ISAFTIEKG 189
QY 251 FGGVTNGKPEDKLGIRGSNTCEVHFENTKIPVENILGEVGDGFKVAMNINLNSGRFSMGV 310
DB 190 FKGFSGIKVEOKLGRASSTTELVELMIVPVENMIGKEGKGFPIAMKTLDGGRIGIAAQ 249
QY 311 VAGLLKRLIEMTAETACTRKOFNKLSEFGLIOEKPALMAQKAYVMSNTYLT--AGMLD 368
DB 250 ALGIAEGAFNEARAYMKERKQGRSLDKF--OGLAMNMADMDVAIESARVLYVKAAYLK 306
QY 369 QGFPDCSTEEAAWVKVFSEAAQVCSEALQILGLGIGYTRDYPPYERILDRTRILLIFGCT 428
DB 307 QAGLP-YTVDAARAKLHAANVAMDVTTRAVQLFGGYGTYTKQYVPERMMRDAKITEIYEGT 365
QY 429 NEILRMVIA 437
DB 366 SEVQKLVIS 374

RESULT 10
ACDP_MYCLE STANDARD; PRT; 389 AA.
AC P46703;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable acyl-CoA dehydrogenase fadE25 (EC 1.3.99.-).
GN FAD25 OR ACD OR ML0737 OR B1308_F1_34.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96084954; PubMed=7489918;
RA Doukhan L., Predlich M., Nair G., Dussurget O., Mandic-Mulec I.,
RA Cole S.T., Smith D.R., Smith I.;
RT "Genomic organization of the mycobacterial sigma gene cluster.";
RL Gene 165:67-70(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltham T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;

```


Db 228 DTCELVFENCPEENVLQVGVVYVLMGLDLRLVLASGPGVIMQACLDVLPYVKQ 287
Qy 329 RKQFNRLSEFGLOEKALMAQKAVVMESMTYLTAGMLDQP--GFPDCSIEAAMVKVFS 386
Db 288 REQFGRIPEFQVQKVDMTYSQSSRSYLYSVARECDSTINTKDC----AGVILSA 343
Qy 387 SEAMQCVSEALQILGGVGYTRDYPYERILRDLTRILLIFEGTNEILRMVY 436
Db 344 AERATQVALQAIQCLGGNGVYNEPTGRFLRDAKLYEAGAGTSEIRRMII 393

RESULT 12
ACDB_HUMAN STANDARD; PRT; 432 AA.
AC P45954; Q96CX7;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Acyl-CoA dehydrogenase, short/branched chain specific, mitochondrial
DE precursor (EC 1.3.99.-) (SBCAD) (2-methyl branched chain acyl-CoA
DE dehydrogenase) (2-MEBCAD) (2-methylbutyryl-coenzyme A dehydrogenase)
DE (2-methylbutyryl-CoA dehydrogenase).
GN ACADSB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95213018; PubMed=7698750;
RA Rozen R., Vockley J., Zhou L., Milos R., Willard J., Fu K.,
RA Vicanek C., Low-Nang L., Torban E., Fournier B.;
RT "Isolation and expression of a cDNA encoding the precursor for a
RT novel member (ACADSB) of the acyl-CoA dehydrogenase gene family";
RL Genomics 24:280-287(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20489853; PubMed=11013134;
RA Andersen B.S., Christensen E., Corydon T.J., Bross P., Pilgaard B.,
RA Wanders R.J.A., Rulter J.P.N., Simonsen H., Winter V., Knudsen I.,
RA Schroeder L.D., Gregersen N., Skovby F.;
RT "Isolated 2-methylbutyrylglutininuria caused by short/branched-chain
RT acyl-CoA dehydrogenase deficiency: identification of a new enzyme
RT defect, resolution of its molecular basis, and evidence for distinct
RT acyl-CoA dehydrogenases in isoleucine and valine metabolism";
RL Am. J. Hum. Genet. 67:1095-1103(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Wambutt R., Heubner D., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Fahy J., Heiton D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Madan A., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP VARIANT 2-METHYLBUTYRYLGLYCINURIA PHE-255.
RX MEDLINE=20290343; PubMed=1083745;
RA Gibson K.M., Burlingame T.G., Hogema B., Jakobs C., Schutgens R.B.H.,
RA Millington D., Roe C.R., Roe D.S., Sweetman L., Steiner R.D.,
RA Linck L., Pohowalla P., Sacks M., Kiss D., Rinaldo P., Vockley J.;
RT "2-methylbutyryl-coenzyme A dehydrogenase deficiency: a new inborn
RT error of L-isoleucine metabolism";
RL Pediatr. Res. 47:830-833(2000).
CC -!- FUNCTION: Has greatest activity toward short branched chain acyl-
CC CoA derivative such as (S)-2-methylbutyryl-CoA, isobutyryl-CoA,
CC and 2-methylhexanoyl-CoA as well as toward short straight chain
CC acyl-CoAs such as butyryl-CoA and hexanoyl-CoA. Can use valproyl-
CC CoA as substrate and may play a role in controlling the metabolic
CC flux of valproic acid in the development of toxicity of this
CC agent.
CC -!- CATALYTIC ACTIVITY: Acyl-CoA + ETF -> 2,3-dehydroacyl-CoA + reduced
CC ETF.
CC -!- COFACTOR: FAD.
CC -!- PATHWAY: Mitochondrial fatty acid beta-oxidation system; first
CC step.
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- DISEASE: Defects in ACADSB are the cause of 2-
CC methylbutyrylglutininuria [MIM:600301]. This is a defect in
CC isoleucine catabolism which is characterized by an increase of 2-
CC methylbutyrylglycine and 2-methylbutyrylcarnitine in blood and
CC urine.
CC -!- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U12778; AAA74424.1; -
DR EMBL; AF260678; AAF97921.1; JOINED.
DR EMBL; AF260668; AAF97921.1; JOINED.
DR EMBL; AF260669; AAF97921.1; JOINED.
DR EMBL; AF260670; AAF97921.1; JOINED.
DR EMBL; AF260671; AAF97921.1; JOINED.
DR EMBL; AF260672; AAF97921.1; JOINED.
DR EMBL; AF260673; AAF97921.1; JOINED.
DR EMBL; AF260674; AAF97921.1; JOINED.
DR EMBL; AF260675; AAF97921.1; JOINED.
DR EMBL; AF260676; AAF97921.1; JOINED.
DR EMBL; AF260677; AAF97921.1; JOINED.
DR EMBL; AL831821; CAD38535.1; -
DR EMBL; BC013756; AAH13756.1; -
DR PIR; A55680; A55680.
DR HSSP; Q06319; 1BUC.
DR Genew; HGNC:91; ACADSB.
DR GK; P45954; -
DR MIM; 600301; -
DR GO; GO:0005739; C:mitochondrion; TAS.
DR GO; GO:0003995; F:acyl-CoA dehydrogenase activity; TAS.
DR GO; GO:0006631; P:fatty acid metabolism; TAS.
DR InterPro; IPR006089; Acyl-CoA_dh.
DR InterPro; IPR006090; Acyl-CoA_dh.
DR InterPro; IPR006091; Acyl-CoA_dh.
DR Pfam; PF00441; Acyl-CoA_dh; 1.
DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
DR Oxidoreductase; Flavoprotein; FAD; Fatty acid metabolism;

KW

KW Mitochondrion; Transit peptide; Disease mutation; Polymorphism.
FT TRANSIT 1 33 MITOCHONDRION.
FT CHAIN 34 432 ACYL-COA DEHYDROGENASE, SHORT/BRANCHED
FT CHAIN SPECIFIC.
FT ACT SITE 414 414 BASE (BY SIMILARITY).
FT VARIANT 209 209 S -> G (in dbSNP:1799823).
FT FTid=VAR_014749.
FT VARIANT 255 255 L -> F (in 2-methylbutyrylglycinuria).
FT FTid=VAR_013010.
FT CONFLICT 13 13 R -> K (IN REF. 3 AND 4).
FT CONFLICT 316 316 I -> V (IN REF. 3).
FT SEQUENCE 432 AA; 47485 MW; 1EB5F894B1944E99 CRC64;
SQ

Query Match 19.28; Score 604; DB 1; Length 432;
Best Local Similarity 35.98; Pred. No. 1.4e-32;
Matches 157; Conservative 81; Mismatches 155; Indels 44; Gaps 14;

Qy 19 GLVSTANRRLLRSTPPVRAFAKELFLGKIKKKEVFP-FPEVSQDE--LNEINQ--FLGP 73
Db 3 GLAV-----RLLRGRLLR-----RNFLTCLSSWKIPPHVSKSSQSQSEALLNTNNGIHFP 53

Qy 74 VEKFFTEVDSRK-----IDQGGKIPDELEKLSGLFLGLQVPEYGG 117
Db 54 LOTETDEEMIKSSVKKFAOQIAPLVSTMDENSKMEKSVIQGLFQOGLMGIEVDPEYGG 113

Qy 118 LG--FSNTMYSLRGELISMDSITVTLAAHQAGLKGILLAGTTEQAKYLPKLASGEHI 175
Db 114 TGASPLTVLV-IEELAKVASVAFCEIQNTLNTLRKGTGEQKATYLPQLTT-EKV 171

Qy 176 AAFCLTEPASGSDAASIRSATLSEDKHYILNGSKVWITNGINGLANITFTPAKTEVVDSD 235
Db 172 GSFCLEAGAGSDSFALKTRADKEGD--YYVLNGSKMWSIAEHAGFLVMANY-----D 224

Qy 236 GSVKDK-ITATIVERDFGVNGRPEDKLGIRGNTCEVHFNKIPVENILGVBGDGFK 294
Db 225 PTIGYKGTISFLVDRDTPGLHIGKPKENKLGURASSTCPLTFENVKVPANILGQIGHGYK 284

Qy 295 VAMNINLGRFSMGVWAGLLKRLTJEMTAETACTKQFNKRLSEFGLIQEKFAKMAQKAY 354
Db 285 YAIQSINTEGRIGIAQMLGLAGCGFDYTPVYKRIQFGKRLDFDQGLQHQVAHVATQLE 344

Qy 355 VMESMTYLTAGMDQPGPDCSIEAMVKVPSSEAAQCVSEALQILGLGTYTRDYPYER 414
Db 345 AARLLTYNAARLL-EAGKPFYK-EASMAKYAYASEIAGATTCKIEWMGGVGYTKDYPYER 402

Qy 415 ILRDTRELLIPETNEI 431
Db 403 YFRDAKIGTYEGASNI 419

RESULT 13
ACDS_MEGEL STANDARD; PRT; 383 AA.
AC Q06319;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acyl-CoA dehydrogenase, short-chain specific (EC 1.3.99.2)
DE (SCAD) (Butyryl-CoA dehydrogenase) (BCAD).
OS Megaspheara elsdeni.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Acidaminococcaceae;
OC Megaspheara.
OX NCBI_Taxid=907;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-50.
RX MEDLINE=94002067; PubMed=8399220;
RA Becker D.F., Fuchs J.A., Banfield D.K., Funk W.D.,
RA Macgillivray R.T.A., Stankovich M.T.;
RT "Characterization of wild-type and an active-site mutant in
RT Escherichia coli of short-chain acyl-CoA dehydrogenase from
RT Megaspheara elsdeni";
RL Biochemistry 32:10736-10742(1993).
RN [2]

RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=95161388; PubMed=7857927;
RA Djordjevic S., Pace C.P., Stankovich M.T., Kim J.J.P.;
RT "Three-dimensional structure of butyryl-CoA dehydrogenase from
RT Megaspheara elsdeni";
RL Biochemistry 34:2163-2171(1995).
CC -!- FUNCTION: HAS AN OPTIMUM SPECIFICITY FOR 4-CARBON LENGTH FATTY
CC ACYL-COAS.
CC -!- CATALYTIC ACTIVITY: Butanoyl-CoA + ETF = 2-butenoyl-CoA + reduced
CC ETF.
CC -!- COFACTOR: FAD.
CC -!- SUBUNIT: Homotetramer.
CC -!- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; L04528; AAA03594.1; -
CC PDB; 1BUC; 20-APR-95.
CC InterPro; IPR006083; Acyl-CoA_dh.
CC InterPro; IPR006090; Acyl-CoA_dh_C.
CC InterPro; IPR006091; Acyl-CoA_dh_M.
CC InterPro; IPR006092; Acyl-CoA_dh_N.
CC Pfam; PF00441; Acyl-CoA_dh; 1.
CC Pfam; PF02770; Acyl-CoA_dh_M; 1.
CC Pfam; PF02771; Acyl-CoA_dh_N; 1.
CC PROSITE; PS00072; ACYL_COA_DH_1; 1.
CC PROSITE; PS00073; ACYL_COA_DH_2; 1.
CC Oxidoreductase; Flavoprotein; FAD; Fatty acid metabolism;
CC 3D-structure.
CC ACT_SITE 367 367 BASE.
CC MUTAGEN 367 367 E->Q: LOSS OF ACTIVITY.
CC FT STRAND 2 2
CC FT HELIX 7 22
CC FT TURN 23 27
CC FT HELIX 28 34
CC FT TURN 35 35
CC FT HELIX 39 46
CC FT TURN 47 48
CC FT HELIX 50 52
CC FT TURN 53 54
CC FT HELIX 57 59
CC FT TURN 60 60
CC FT HELIX 62 65
CC FT TURN 66 66
CC FT HELIX 69 82
CC FT HELIX 84 96
CC FT TURN 97 97
CC FT HELIX 98 104
CC FT HELIX 107 112
CC FT TURN 113 113
CC FT HELIX 114 119
CC FT TURN 120 120
CC FT STRAND 124 127
CC FT STRAND 130 130
CC FT TURN 131 132
CC FT STRAND 133 133
CC FT HELIX 137 139
CC FT STRAND 143 146
CC FT TURN 148 149
CC FT STRAND 152 162
CC FT TURN 163 166
CC FT STRAND 169 175
CC FT TURN 182 183
CC FT STRAND 185 191
CC FT TURN 192 193
CC FT TURN 195 196
CC FT STRAND 197 202

FT	STRAND	209	210
FT	STRAND	213	223
FT	HELIX	225	227
FT	STRAND	228	229
FT	TURN	232	233
FT	STRAND	234	234
FT	HELIX	236	237
FT	STRAND	238	241
FT	TURN	245	246
FT	STRAND	247	248
FT	STRAND	250	251
FT	STRAND	252	253
FT	STRAND	255	256
FT	STRAND	257	258
FT	HELIX	259	261
FT	HELIX	263	312
FT	TURN	313	313
FT	HELIX	317	342
FT	STRAND	343	347
FT	TURN	349	350
FT	HELIX	353	360
FT	HELIX	361	364
FT	TURN	365	367
FT	HELIX	370	381
FT	TURN	382	383
SQ	SEQUENCE	383 AA; 41408 MW; 3D68AAE34D9BBAB8 CRC64;	

Query Match 18.8%; Score 593.5; DB 1; Length 383;
Best Local Similarity 38.1%; Pred. No. 5,8e-32;
Matches 149; Conservative 67; Mismatches 156; Indels 19; Gaps 9;

QY	54	PFFPEVSQDELNEINQFLGPVEKFTTEEVDSRKIDQEGKIPDETLEKLKSLGLFQLQVPE	113
DB	3	FNLTDIOQDFLKLAHDF---GEKKLAPTPTTER--DHKGIIYDKELIDELLSLGITGAYFEE	57
QY	114	EYGGLGFSGN----TWISRLGRIIISMDGSITVTLLAHOQAIGLKGIILAGTSEQAKYLPKL	169
DB	58	KYGGSGDDGDVLSYILAWEELAKYDAGVAITLSATVSLCANPIWQFGTEAQKEKFLVPL	117
QY	170	ASGHIAAFCLTEPAGSDAASIRSRTLSEDRKHYYILNGSKVWTNGGLANIFTVFATK	229
DB	118	VEGTKLGAFLGTETPNAGTDSAGQQTATKNDDGT-YTLNGSKIITNGRAIIVIVAMT	176
QY	230	EVVSDSGVKDKITAFIVERDFGVTKNGKPEDKLGIRGNSNTCEVHFENTKIPIVENILGEV	289
DB	177	D--KSKGN--HGITAFLIEDGTGFTYGGKEDKMGHIHTSQTMELVFQDVKVPAENMLGEE	232
QY	290	GDFGFKVAMTNLNSGRFSVGMSVAGLLRRLIEMTAEYACTRKQFNKRlsefGLIOEKPALM	349
DB	233	GKGFKIAMMTLDGGRIGCVAAALGIAEAALADAVEYSKQRVQFGKPLCKQSISFKLAM	292
QY	350	AQRAYVNESWTYLTAGMLDQPGPDSCIEAAMVVFSSAEAWQCVCSEALQILGSLGYTRD	409
DB	293	KMOIEARNVLY-KACKKQEGKP-FVTDAAIAKRVASDVMARVTTTEAVQIFGGYGYSEE	350
QY	410	YPYERILDRILLIFEETNEILRMVIALTG 440	
DB	351	YPVARHMRDAKITQIVEGTNEVOLM---VTG 378	

RESULT 14
ID_VIDL_SOLUTE STANDARD; PRT; 412 AA.
AC Q9FS88;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Isovaleryl-CoA dehydrogenase 1, mitochondrial precursor (EC 1.3.99.10) (VID 1).
DE DE
GN IDVI.
OC Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.

RESULT 15
 IVD_HUMAN
 ID IVD_HUMAN STANDARD; PRT; 423 AA.
 AC P26440; Q96AF6;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Isovvaleryl-CoA dehydrogenase, mitochondrial precursor (EC 1.3.99.10)
 DE (IVD).
 GN IVD.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90203210; PubMed=2318964;
 RA Matsubara Y., Ito M., Glassberg R., Satyabhama S., Ikeda Y.,
 RA Tanaka K.;
 RA "Nucleotide sequence of messenger RNA encoding human isovaleryl-
 RA coenzyme A dehydrogenase and its expression in isovaleric acidemia
 RA fibroblasts";
 RL J. Clin. Invest. 85:1058-1064(1990).
 [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20143360; PubMed=10677295;
 RA Vockley J., Rogan P.K., Anderson B.D., Willard J., Seelan R.S.,
 RA Smith D.I., Liu W.;
 RA "Exon skipping in IVD RNA processing in isovaleric acidemia caused by
 RA point mutations in the coding region of the IVD gene";
 RL Am. J. Hum. Genet. 66:356-367(2000).
 [3]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Brain;
 RC MEDLINE=22386257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalski U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length
 RA human and mouse cDNA sequences";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [4]
 RP SEQUENCE OF 381-423 FROM N.A.
 RX MEDLINE=93224146; PubMed=8468053;
 RA Parimoo B., Tanaka K.;
 RA "Structural organization of the human isovaleryl-CoA dehydrogenase
 RA gene";
 RL Genomics 15:582-590(1993).
 [5]
 RP ACTIVE SITE GLU-283.
 RX MEDLINE=95367542; PubMed=7640268;
 RA Mohsen A.W., Vockley J.;
 RA "Identification of the active site catalytic residue in human
 RA isovaleryl-CoA dehydrogenase";
 RL Biochemistry 34:10146-10152(1995).
 [6]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-421.
 RX MEDLINE=97361954; PubMed=9214289;
 RA Tiffany K.A., Roberts D.L., Wang M., Paschke R., Mohsen A.W.,
 RA Vockley J., Kim J.J.;
 RA "Structure of human isovaleryl-CoA dehydrogenase at 2.6-A resolution:
 RA structural basis for substrate specificity";
 RL Biochemistry 36:8455-8464(1997).
 [7]
 RP VARIANTS IVA PRO-42 AND VAL-199.
 RX MEDLINE=91290048; PubMed=2063866;
 RA Vockley J., Parimoo B., Tanaka K.;
 RA "Molecular characterization of four different classes of mutations in
 RA the isovaleryl-CoA dehydrogenase gene responsible for isovaleric
 RA acidemia";
 RL Am. J. Hum. Genet. 49:147-157(1991).
 CC -!- CATALYTIC ACTIVITY: 3-methylbutanoyl-CoA + ETF = 3-methylbut-2-
 CC enoyl-CoA + reduced ETF.
 CC -!- COFACTOR: FAD.
 CC -!- PATHWAY: Leucine catabolism.
 CC -!- SUBUNIT: Homotetramer.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -!- DISEASE: Defects in IVD are the cause of isovaleric acidemia (IVA)
 CC [MIM:243500]. IVA is characterized by retarded psychomotor
 CC development, a peculiar odor resembling sweaty feet, an aversion
 CC to dietary protein, and pernicious vomiting, leading to acidosis
 CC and coma. The acute neonatal form leads to massive metabolic
 CC acidosis from the first days of life and rapid death.
 CC -!- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M34192; AAA52711.1; -
 CC EMBL; AF191218; AAF20182.1; JOINED.
 CC EMBL; AF191214; AAF20182.1; JOINED.
 CC EMBL; AF191215; AAF20182.1; JOINED.
 CC EMBL; AF191216; AAF20182.1; JOINED.
 CC EMBL; AF191217; AAF20182.1; JOINED.
 CC EMBL; BC017202; AAH17202.1; -
 CC EMBL; AF038318; AAB92584.1; -
 CC PIR; A37033; A37033.
 CC PDB; 1IVH; 20-MAY-98.
 CC Genew: HGNC:6186; IVD.
 CC GK; P26440; -
 CC MIM; 243500; -
 CC GO; GO:0008470; F:isovaleryl-CoA dehydrogenase activity; TAS.
 CC InterPro; IPR006089; Acyl-CoA_dh.
 CC InterPro; IPR006090; Acyl-CoA_dh_c.
 CC InterPro; IPR006091; Acyl-CoA_dh_m.
 CC InterPro; IPR006092; Acyl-CoA_dh_n.
 CC Pfam; PF00441; Acyl-CoA_dh; 1.
 CC Pfam; PF02770; Acyl-CoA_dh_m; 1.
 CC Pfam; PF02771; Acyl-CoA_dh_n; 1.
 CC PROSITE; PS00072; ACYL_COA_DH_1; 1.
 CC PROSITE; PS00073; ACYL_COA_DH_2; 1.
 CC Oxidoreductase; Flavoprotein; FAD; Mitochondrion; Transit peptide;
 KW Disease mutation; 3D-structure.
 KW TRANSIT
 FT CHAIN 1 29 MITOCHONDRION
 FT CHAIN 30 423 ISOVALERYL-COA DEHYDROGENASE.
 FT ACT_SITE 283 283
 FT VARIANT 42 42
 FT
 FT VARIANT 199 199
 FT
 FT MUTAGEN 283 283 L -> P (in IVA).
 FT MUTAGEN 283 283 G -> V (in IVA).
 FT CONFLICT 10 10 /FTID=VAR_000423.
 FT HELIX 38 40 /FTID=VAR_000424.
 FT TURN 41 41 E->G,Q: LOSS OF ACTIVITY.
 FT HELIX 41 41 W -> C (IN REF. 3).
 FT TURN 44 44
 FT HELIX 44 44

FT TURN 61 64
FT HELIX 65 71
FT TURN 72 72
FT TURN 75 76
FT HELIX 77 87
FT TURN 88 88
FT TURN 90 93
FT HELIX 96 98
FT TURN 99 99
FT HELIX 105 118
FT HELIX 120 130
FT TURN 131 133
FT HELIX 134 140
FT HELIX 143 155
FT TURN 156 156
FT STRAND 160 163
FT STRAND 166 166
FT TURN 167 168
FT STRAND 169 169
FT HELIX 173 175
FT STRAND 179 182
FT STRAND 186 197
FT TURN 198 198
FT HELIX 199 201
FT STRAND 204 210
FT TURN 213 214
FT HELIX 218 221
FT STRAND 222 228
FT TURN 229 230
FT TURN 232 233
FT STRAND 234 236
FT STRAND 242 242
FT TURN 246 247
FT STRAND 250 261
FT HELIX 262 264
FT STRAND 265 266
FT TURN 269 270
FT HELIX 272 287
FT TURN 288 288
FT HELIX 289 306
FT TURN 307 308
FT STRAND 310 311
FT TURN 312 313
FT STRAND 314 315
FT HELIX 316 318
FT TURN 320 348
FT TURN 349 350
FT HELIX 354 379
FT STRAND 380 384
FT TURN 385 385
FT STRAND 386 387
FT TURN 389 389
FT HELIX 390 399
FT TURN 400 404
FT HELIX 407 419
FT TURN 420 420

Query Match 18.6%; Score 587.5; DB 1; Length 423;
Best Local Similarity 34.1%; Pred. No. 1.7e-31;
Matches 154; Conservative 78; Mismatches 172; Indels 47; Gaps 13;
QY 22 VSTANRL-----LRTSPVRAFAKELFLGKIKKKEVFPFPEVSQDELNEINQFLGPVE 75
DB 1 MATATRLGWRVAVRRLPPLAGEVSQ-----RAHSLLPVDDAINGLSEQRQLRQTMA 54
QY 76 KFTTEV--DSRKIDQCKIPD--ETLEKLSGLGLQVPPEYGGGLGFSNTMYSRLGEI 131
DB 55 KFLQEHAPKAQEIERSNEFNKLNREFWKQLGNLGLVGTAPVQYGGSLGLYLEHVLVWEE 114
QY 132 IS-MDGSITVTLAHOAIGLGIILAGTEQAKYLPKLASGEHIAAFLTEPASGSDAA 190
DB 115 ISRASGAVGLSYGAHSLCINQLVRNGNEAQEKYLPKLISGEYIGALANSEPNAGSDVV 174

QY 191 SIRSRATLSEDKK--HYILNGSKVWITNGGLANIETVFAKTEVVDSDGSVKDKITAFIVE 248
DB 175 SMKLKA-----EKKGNHYILNGNKFVITNGPDADVLIVYAKTDLAAVPAS--RGITAFIVE 228
QY 249 RDFGGVTNGKPEDKLGIRGNTCEVHFENTKIPVENILGEVGDGFKVAMNILNSGRFSMG 308
DB 229 KCMPGFSTSKKLDKLGMRGNTCELFEDCKIPAAANILCHENKGVYVLMGSLDLRLVLA 288
QY 309 SVVAGLLKRLIEMTAETAYACTRKQFNKRLSEFGLIOEKFA-----LMAQKAYVMESMTYLT 363
DB 289 GGPLGLMOAVLDHTIPYLHVREAFQKIGHFOLMOGKNADMYTRLMACRQYV-----YNV 343
QY 364 AGMLDQPGPPDCSI-FAAMVKVFSSEAAQCVSEALQILGGLGYTRDYPYERILDRIL 422
DB 344 AKACDE---GHCTAKDCAGVILYSAECATQVALDGIQCGGNGYINDFPMGREFLRDAKLY 400
QY 423 LIFEETNEILRMVIALTGLQHAGRIITRIH 453
DB 401 EIGACTSEVRRLVI-----GRAFNAFDH 423

Search completed: September 3, 2003, 13:28:03
Job time : 27 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 13:24:57; Search time 42 seconds
(without alignments)
1421.923 Million cell updates/sec

Title: US-09-945-326-2
Perfect score: 3153
Sequence: 1 MSGCGLFLRTTAAARACRGL.....SQQLEKRAYICAHPLDRTC 621

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR.76.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3149	99.9	621	2 JC7892	acyl-CoA dehydroge
2	1356.5	43.0	655	2 A54872	acyl-CoA dehydroge
3	1308	41.5	655	2 S54183	acyl-CoA dehydroge
4	1210.5	38.4	613	2 T15905	hypothetical prote
5	821	26.0	594	2 F84085	butyryl-CoA dehydr
6	815.5	25.9	594	2 C70021	butyryl-CoA dehydr
7	711	22.5	650	2 A70817	probable fadE10 pr
8	699	22.2	382	2 B87302	acyl-CoA dehydroge
9	691.5	21.9	384	2 B83443	acyl-CoA dehydroge
10	690	21.9	379	2 G84124	acyl-CoA dehydroge
11	682	21.6	378	2 D69658	acyl-CoA dehydroge
12	682	21.6	381	2 AC3536	butyryl-CoA dehydr
13	678	21.5	381	2 B87472	acyl-CoA dehydroge
14	658	20.9	380	2 F84124	acyl-CoA dehydroge
15	652.5	20.7	379	1 S55421	acyl-CoA dehydroge
16	652	20.7	387	2 D75417	acyl-CoA dehydroge
17	648	20.6	379	2 T45286	butyryl-CoA dehydr
18	630	20.0	389	2 C70979	probable fadE25 pr
19	629.5	20.0	381	2 E75458	acyl-CoA dehydroge
20	628	19.9	456	2 AF3637	butyryl-CoA dehydr
21	626	19.9	380	2 A12983	acyl-CoA dehydroge
22	626	19.9	380	2 F98299	probable acyl-CoA
23	625.5	19.8	379	2 T47262	butyryl-CoA dehydr
24	625.5	19.8	379	2 F97233	butyryl-CoA dehydr
25	617.5	19.6	387	2 A83393	probable acyl-CoA
26	615	19.5	419	2 T15088	hypothetical prote
27	613	19.4	380	2 T44811	acyl-CoA dehydroge
28	611.5	19.4	402	2 D69530	probable acyl-CoA
29	610.5	19.4	381	2 B83791	butyryl-CoA dehydr

ALIGNMENTS

RESULT 1

JC7892

acyl-CoA dehydrogenase (EC 1.3.99.3) - 9 - human

C:Species: Homo sapiens (man)

C:Date: 03-Feb-2003 #sequence_revision 03-Feb-2003 #text_change 14-Apr-2003

C:Accession: JC7892

R:Zhang, J., Zhang, W.; Zou, D.; Chen, G.; Wan, T.; Zhang, M.; Cao, X.

Biochem Biophys Res Commun 297:103-104, 2002

A:Title: Cloning and functional characterization of hACAD9, a novel member of human a

-long-chain acyl-CoA and has specific activity toward palmitoyl-CoA (C16:0).

A:Reference number: JC7892; MUID:22246103; PMID:12359260

A:Accession: JC7892

A:Molecule type: mRNA

A:Residues: 1-621 <ZHA>

A:Cross-references: GB:AF327351

A:Experimental source: Dendritic cells

C:Comment: This enzyme, which belongs to a family of mitochondrial enzymes that catal

-long-chain acyl-CoA and has specific activity toward palmitoyl-CoA (C16:0).

C:Genetics:

A:Gene: acad-9

A:Map position: 3q26

C:Keywords: oxidoreductase

Query Match 99.9%; Score 3149; DB 2; Length 621;

Best Local Similarity 99.8%; Pred. No. 6.6e-201;

Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MSGCGLFLRTTAAARACRGLVSTANRRLLRTSPVPRAFPAKELFLGKIKKKEVFPPEVS	60
DB	1	MSGCGLFLRTTAAARACRGLVSTANRRLLRTSPVPRAFPAKELFLGKIKKKEVFPPEVS	60
QY	61	QDELNEINFLGPVEKFFTEEVDSRKIDQEKIPDETLEKLSGLFLGLQVPEEYGGGLF	120
DB	61	QDELNEINFLGPVEKFFTEEVDSRKIDQEKIPDETLEKLSGLFLGLQVPEEYGGGLF	120
QY	121	SNTMYSRLGEIISMDGSIITVTLAAHQAGLKGIIILAGTEEQAKYLPKLASGEHIAAFCL	180
DB	121	SNTMYSRLGEIISMDGSIITVTLAAHQAGLKGIIILAGTEEQAKYLPKLASGEHIAAFCL	180
QY	181	TEPASGSDAASRSRATLSEDKKHKLNGSKVWITNGGLANIFTVFAKTEVVDSDGSKVD	240
DB	181	TEPASGSDAASRSRATLSEDKKHKLNGSKVWITNGGLANIFTVFAKTEVVDSDGSKVD	240
QY	241	KITATIVERDFGVTNGKPEDKLGIRGNTCVHFNTPKIPVENTLGEVGGFKVAMNLL	300
DB	241	KITATIVERDFGVTNGKPEDKLGIRGNTCVHFNTPKIPVENTLGEVGGFKVAMNLL	300
QY	301	NSGRFSMGSVVAGLLKRLIEMTAETACTKQFNKRLSEFGLTQERFALMAQAKAYVMSMT	360
DB	301	NSGRFSMGSVVAGLLKRLIEMTAETACTKQFNKRLSEFGLTQERFALMAQAKAYVMSMT	360
QY	361	YLTAGMLDQGFPPDCSIEAAMVKVFPSSAANCCVSEALQILGGGLGYTRDYPYERILDR	420
DB	361	YLTAGMLDQGFPPDCSIEAAMVKVFPSSAANCCVSEALQILGGGLGYTRDYPYERILDR	420

not nt

C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C:Accession: C70021
R:Kunst, F.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Broutlet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Ferrar, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
lech, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
y, M.; Ogawa, K.; Ogata, B.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tonon, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: C70021
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-594 <KUN>
A:Cross-references: GB:299120; GB:AL009126; NID:g2635613; PIDN:CAB15271.1; PID:ell184360;
A:Experimental source: strain 168
C:Genetics:
A:Gene: yusJ

Query Match 25.9%; Score 815.5; DB 2; Length 594;
Best Local Similarity 33.4%; Pred. No. 3e-46;
Matches 201; Conservative 97; Mismatches 202; Indels 101; Gaps 12;
QY 59 VSQDELNEINFLGPEVKEFFTEEDSRKIDQEGKIPDETLEKLSGLGLQVPEYGG 118
DB 45 IEQDVLPPI-----DDIENHQFHSVRL-----LKKAGELGLGADVPEYGG 88
QY 119 GFSNTMYSLRGLIISMGDSITVTLAAHQATGLKGIILAGTEPEQAKYLPKLASEHTAAF 178
DB 89 GLDKISSALITKFSRAGSFLSFGAHVIGISLPIVFGESEQKKYLPGLGASEKTAAY 148
QY 179 CLTPASGSDAASRSTRATLSSEDKKHYTLNLSKVWITNGGLIANTFTVFAKTEVVDSDGS 238
DB 149 ALTEPGSSDALGAKTTAVLNEAGTHVLTGEKWIINSADVFVYAKV-----DG-- 201
QY 239 KDKITAFIVERDFGVTNGKPEDKLIGRNSNTCEVHENTKIPVENILGEVGDGFKVAMN 298
DB 202 -DKTSAFIVEKFCVGTGPEKKMGKIGSSTRLLILDQAEVPRKENVLLGEGKHVIAFN 260
QY 299 ILNSGRFSGVWAGLLKRLIEMTAETACTKQFNKRLSEFLGLOEKFAALMAQKAYVNES 358
DB 261 ILNIGRYKLAVGTIGASKRVIELSAAYANQRQKFTPIAGPSLTQEKIGTWASRLYAMES 320
QY 359 MTLTAGMLDQ-----PGPDCSIEAAMVKVFSSEAAWQCVSEALQIL 401
DB 321 SVYRTVGLFEDNMSQFTAECLKDGRQIAKSAEVAIECSLNKVFGESETLDIVDEGVQIH 380
QY 402 GGLGYTRYDYPYERILDRTRILLIPGTNEILRMVYALTGLQHA--GRI-LITRHELKQA 458
DB 381 GGYGFMQYEVBRAYRDSRINRIFEGTNEINRLIVPSTFLKALKGELPULPEKAQSLQEE 440
QY 459 KYSTVMDTVGRRLDSGLRTVDLGTGNHGVHPSPADSANKFENTCYFCGRTVETLL-- 517
DB 441 LMKLMEPEFG-----SGVL-----EKEKIVVQAKKIALFA 471
QY 518 -----RFKGTIMEQLVLRVANILINLYGMTAVLSRASRSIRI-GLRNHDEHVLANT 570
DB 472 AGLAAQKYGKADREQIEILVNVIVSNVYAMESAVLRTEKAIAQAQKAKQKLYLTYE 531
QY 571 FCVEAYLO-----NLSLSOLDKYAPNLDQIKKVSQQLLEKRA 610
DB 532 FVOEAFNEIEAHAKESLIAMEGDSLRMWLSALRKLTRVTPKNVIQKKREAAGIFEAK 591
QY 611 Y 611

Db 592 Y 592

RESULT 7

A70817
probable fadE10 protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: A70817
R:Colet, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
C:Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: A70817
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-650 <COL>
A:Cross-references: GB:AL022004; GB:AL123456; NID:g3261550; PIDN:CAAL17679.1; PID:el129
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: fadE10

Query Match 22.5%; Score 711; DB 2; Length 650;
Best Local Similarity 31.4%; Pred. No. 2.9e-39;
Matches 194; Conservative 113; Mismatches 223; Indels 88; Gaps 16;
QY 38 AFAKELFLGKIKKKEVFPPEVSDQELNEINQFLGPVKEKFFTEEDSRKIDQEGKIPDET 97
DB 28 SFAKELFLGPRPLGLIHPPKPSDAEAEAFVLRKREFL-DTVDGVSIVIERAAQIPDEY 86
QY 98 LEXLSLGLRGLQVPEYGGIFGFSNTMYSR-LGEIISMDGSIIVTLAAHQATGL-KGIIL 155
DB 87 VKGLAELGCGFLGKIPSEYGGNLNMSQVAYNRVLMVTVTHSSLSGALLSAHOSIGVPEPLK 146
QY 156 AGTEQAKYLPKLASEHTAAFLTPASGSDAASRSTRATLSSEDKKHYLNGSKVWIT 215
DB 147 AGTAEOKRRFLPCAAGA-ISAFLLTEPDVGDSPARMASATPIDDQQAQYELGKLVTT 205
QY 216 NGGLANITFTVFAKTEVVDSDGSVKDKITAFIVERDFGVTNGKPEDKLIGRNSNTCEVHF 275
DB 206 NGVVADLLVMAR--VPRSEGH-RGGISAFVVEADSPGITVERNRKPMGLRGIENGVTRL 262
QY 276 ENTKIPVENILGEVGDGFKVAMNILNSGRFSMGVAVAGLLKRLIEMTAETACTKQFNKR 335
DB 263 HRVVPKDNLTIGREGDGLKIALTLNAGRLSLPAIATGVAKQALUKIAREWSVERVQMGKP 322
QY 336 LSEFGLIOEKFAALMAQKAYVNESMTYLTAGLDQPGPDCSIEAAMVKVFSSEAAWQCVS 395
DB 323 VGQHEAVASKISFIAATNYALDAVVELSSQMADE-GRNDFIEAALAKINSEMAVLGD 381
QY 396 EALQILGLGLY-----TRDYPYERILDRTRILLIPGTNEILRMVYALTGLG----L 441
DB 382 ELQIRGGRGYETAESLAARGERAVPVQMVYRDLRINRIFEGSSEIMRLIAREAVDAHL 441
QY 442 QHAGRLITRHELKQAKVSTVMDTVGRRLDSGLRTVDLGTGNHGVHPSPADSANKFENTCYFCGRTVETLL-- 495
DB 442 TAAGDLNPKA-DLQKAAAAA-----GASGYAKWLPKLVFGEQOL 482
QY 496 -----DSANKFENTCYFCGRTVETLLLRFGKTIEMBEQVLKRVANIL 537
DB 483 PTTYREFCALATHLRFVRSRKLARNTF-YG-----MARWQASLEKKQKQGLGRIVDIG 535
QY 538 INLYGMTAVLSRASRSIRI-GLRNHDEHVLANTFC-----VEAYLONLFSLSOLDKYAP 591
DB 536 AELFAISACVRAEQRTADPVEGEQAYELAEAFCCQOATLRVEALFDALWS----- 586
QY 592 ENLDEQIKKVSQQLLEKR 609
DB 587 -NTSDIVRLANDVLEGR 603

RESULT 8

B87302
acyl-CoA dehydrogenase, probable [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: B87302
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete genome sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: B87302
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-382 <STO>
A:Cross-references: GB:AE005673; NID:gl3421592; PIDN:AAK22414.1; GSPDB:GN00148
C:Genetics:
C:Superfamily: acyl-CoA dehydrogenase

Query Match 22.2%; Score 599; DB 2; Length 382;
Best Local Similarity 38.5%; Pred. No. 8.2e-39;
Matches 155; Conservative 80; Mismatches 124; Indels 44; Gaps 8;
Qy 63 ELNEINQFLGPKVEKFFTEVD--SRKIDQEKIPDETLEKLSGLGFLQVPEEYGGGLG- 120
Db 4 DLETREQLDITVAREFAERLPIEAQVAENDAVDDVIEEMKGLGLGLTIPEEFGGLG- 62
Qy 121 SNTWSRGLGELISMDGSIITVTLAAHQ-----IGLKGIILAGTEEQAKYLP 167
Db 63 -----LTMEERALVAIELGRASPAPRSVFGTNGVIGSGGLWFGNDEQAKWLP 111
Qy 168 KLASGEHIAFLTEPAGSDAASIRSRATLSEDKKHYILNGSKVWTNGLANIFTVFA 227
Db 112 GIASGAVITSFALTEPAGSDSAVQTRAT--RDGDDYILNGSKRYITNAGKASLETVMA 169
Qy 228 KTEVVDSDGVKDKITAFIVERDGGVTNGKPKDLGIRGNTCEVHFENTKIPVENILG 287
Db 170 RTN-PDAKGA--GVSAFLVPRDLPGITVGKPKKMGQGAHHDVTFDNRVFPAMNRLG 226
Qy 288 EVGDGFKVAMNILNGSRFSGVAVGLLKRLEIEMTAETACTRKQFNKRLSEFGLIQ--- 343
Db 227 AEGGFKVAMQVLDGRGLHIAACVGAERLIADCVAYASERKQFGQPIASFOLIQAIA 286
Qy 344 -EKFALMAOKAYVMESMTYLTAGMLDQPGFPCDSIEAAMVKVFSSEAAMQCVSEALQILG 402
Db 287 DSTELAAALVLETARKRDAGV-----NVTLEAASKLPFASEMVGVRADRAVQVFG 339
Qy 403 GLGYTRDYPYERILDRTRILLIFEGTNEILRMVIALTGLQHAG 445
Db 340 GAGVADYGIERLYRDVRIEYEGTSQVQQLIARETLKRG 382

RESULT 9

B83443
probable acyl-CoA dehydrogenase PA1631 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83443
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Llin
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: B83443
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-384 <STO>
A:Cross-references: GB:AE004590; GB:AE004091; NID:99947587; PIDN:AAG05020.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA1631

C:Superfamily: acyl-CoA dehydrogenase

Query Match 21.9%; Score 691.5; DB 2; Length 384;
Best Local Similarity 37.8%; Pred. No. 2.6e-38;
Matches 147; Conservative 82; Mismatches 121; Indels 39; Gaps 8;
Qy 67 INQFLGPKVEKFFTEV--DSRKIDQEKIPDETLEKLSGLGFLQVPEEYGGGLGFSNTM 124
Db 8 LNLLEDISIRQFVRESLVPHEQVAETDRIPEAIIARMREMGFLGSLSPAYGGGLG----- 62
Qy 125 YSRGLGELISMDGSTTIVT-----LAAHQALGLKGIILAGTEEQAKYLPKLAS 171
Db 63 -----VTMEEEVSIAPFELGRTPAPRSLTLGTTNNGIGSQGIIVDTGEQKRRYLPRLAS 115
Qy 172 GEHIAAFCLTEPAGSDAASIRSRATLSEDKKHYILNGSKVWTNGLANIFTVFAKTEV 231
Db 116 GELLSSFLTEPDSGSDAASLKTAV--RGEHYVLNGTKRFTTNAPQAGIYTWART-- 171
Qy 232 VSDGSVKDK--ITAFIVERDGGVTNGKPKDLGIRGNTCEVHFENTKIPVENILGEV 289
Db 172 ---DPAIRGAGGISAFVVERGTPLGSLGKPKDRKMGKGAHTCDVIFDCRPASQLIGV 228
Qy 290 -GDGFKVAMNILNGSRFSGVAVGLLKRLEIEMTAETACTRKQFNKRLSEFGLIOEKFAL 348
Db 229 EGVGFKTAMVKLOKRLHIAACVGAERLMDALRYALERKQFGQPIAEFQLIQAMLAD 288
Qy 349 MAQKAYVMESMTYLTAGMLDQPGFPCDSIEAAMVKVFSSEAAMQCVSEALQILGGLGYTR 408
Db 289 SKAEVAAARCMVIDAARQRDEG--RDVGTASCAKLPFASEMVGVRADRAVQIFGGAGYIG 346
Qy 409 DYPYERILDRTRILLIFEGTNEILRMVIA 437
Db 347 DYGIERYDRVRLFRIVEGTTIQQLLIA 375

RESULT 10

G84124
acyl-CoA dehydrogenase mmgC [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: G84124
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: G84124
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-379 <STO>
A:Cross-references: GB:AP001520; GB:BA000004; NID:gl0176401; PIDN:BAB07518.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
C:Gene: mmgC
C:Superfamily: acyl-CoA dehydrogenase

Query Match 21.9%; Score 690; DB 2; Length 379;
Best Local Similarity 40.8%; Pred. No. 3.2e-38;
Matches 147; Conservative 74; Mismatches 125; Indels 14; Gaps 6;
Qy 78 FTEEVDSRKIDQEGKIPDETLEKLSGLGFLQVPEEYGGGLGFSNTMY-SRLGEIISMDG 136
Db 27 FVEEMET-----NDTFPMHIVKEMGELGMLGPIPEAYGGAEMDMFTSYTIIAHELSKVS 81
Qy 137 SITVTTLAAHQALGLKGIILAGTEEQAKYLPKLASGEHIAAFCLTEPAGSDAASIRSR 196
Db 82 TVGVILSVHTSVGTNPILFEGTEEQSKRIPKLAGEYLKAGLTPSAGSDAANLKTFA 141
Qy 197 TLSEDKKHYILNGSKVWTNGLANIFTVFAKTEVVDSDGSVKDKITAFIVERDGGVTN 256
Db 142 LKQGD--HYRLNGSKVFTINGGAADTVVFAST----NPSAGRKGISAFIVEKETPGFTV 195

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2003, 10:14:14 ; Search time 6852.73 Seconds
(without alignments)
11121.789 Million cell updates/sec

Title: US-09-945-326-3
Perfect score: 1863
Sequence: 1 atgagcgtcggtggtctt.....accctctggacaggacatgc 1863

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_on:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_to:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1863	100.0	1863	6	AX394749	AX394749 Sequence
2	1863	100.0	2448	9	BC013354	BC013354 Homo sapi
3	1863	100.0	2452	6	AX394747	AX394747 Sequence
4	1859.8	99.8	2399	6	BD158773	BD158773 Primer fo
5	1859.8	99.8	2399	9	AK024012	AK024012 Homo sapi
6	1859.8	99.8	2494	9	BC007970	BC007970 Homo sapi
7	1858.2	99.7	2411	9	AF327351	AF327351 Homo sapi
8	1671.4	89.7	2400	6	BD157796	BD157796 Primer fo
9	1671.4	89.7	2400	9	AK022568	AK022568 Homo sapi
10	1380.8	74.1	2129	10	BC033277	BC033277 Mus muscu
11	1380.8	74.1	2148	10	BC032213	BC032213 Mus muscu
12	1380.8	74.1	3071	10	BC031137	BC031137 Mus muscu
13	1029	55.2	1494	9	BC001817	BC001817 Homo sapi
14	704.6	37.8	771	6	BD150623	BD150623 Primer fo
15	526	28.2	1995	9	AK027186	AK027186 Homo sapi
16	499.6	26.8	772	6	BD129328	BD129328 Human gen
17	496.4	26.6	544	6	BD109621	BD109621 EST and e
18	484.4	26.0	543	6	AX588711	AX588711 Sequence
19	460.4	24.7	507	6	AX588841	AX588841 Sequence
20	418.4	22.5	431	6	AX072562	AX072562 Sequence
21	396.4	21.3	2058	9	AB083302	AB083302 Macaca fa
22	396.4	21.3	2103	10	MTMVLCAD	271189 M.musculus
23	396	21.3	2140	4	BTU30817	U30817 Bos taurus
24	394.8	21.2	2171	10	BC026559	BC026559 Mus muscu
25	393.2	21.1	796	6	BD147321	BD147321 Primer fo
26	391.6	21.0	1965	6	BD141736	BD141736 IRAP bind
27	391.6	21.0	2168	9	BC020218	BC020218 Homo sapi
28	391.6	21.0	2169	9	HSHVLCAD	D43682 Human mRNA
29	391.6	21.0	2219	9	HSHVLCAD	X86556 H.sapiens H
30	391.6	21.0	2231	9	BC012912	BC012912 Homo sapi
31	391.6	21.0	2258	9	AK056083	AK056083 Homo sapi
32	391.6	21.0	2259	9	BC000399	BC000399 Homo sapi
33	391.2	21.0	1840	10	AF017176	AF017176 Mus muscu
34	388.6	20.9	2117	6	AX401908	AX401908 Sequence
35	388.6	20.9	2117	10	RATVLCAD	D30647 Rat mRNA fo
36	383	20.6	1439	10	AY292988	AY292988 Rattus no
37	380.2	20.4	434	6	AX071059	AX071059 Sequence
38	372.8	20.0	171981	3	AC008344	AC008344 Drosophil
39	372.8	20.0	176813	3	AC008096	AC008096 Drosophil
40	372.8	20.0	176980	2	AC020290	AC020290 Drosophil
41	372.8	20.0	277421	3	AE003796	AE003796 Drosophil
42	372.6	20.0	412	6	AX071159	AX071159 Sequence
43	339.8	18.2	2027	3	AK115332	AK115332 Ciona int
44	283.2	15.2	300	6	BD129040	BD129040 Human gen
45	283	15.2	1476	9	AK058109	AK058109 Homo sapi

ALIGNMENTS

RESULT 1
AX394749
LOCUS AX394749 1863 bp DNA linear PAT 18-MAY-2002
DEFINITION Sequence 3 from Patent WO0218582.
ACCESSION AX394749
VERSION AX394749.1 GI:21065828
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Meyers, R. and Hunter, J.J.
TITLE 62112, a novel human dehydrogenase and uses thereof
JOURNAL Patent: WO 0218582-A 3 07-MAR-2002;

Millennium Pharmaceuticals, Inc. (US)		
FEATURES	Location/Qualifiers	
source	1. .1863	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
CDS	1. .>1863	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAD32050.1"	
	/db_xref="GI:21065829"	
	/translation="MSGGFLRTTAAARACRLVSTANRLLRTSPVRAFAPLELF	
	LGRIKKKEVFPFVQSGDELNEINQFVPEFFTEEDSRKIDQEGKIPDETLEKLK	
	SIGLFGLOVPEYGGGFSNTMYSLRGEIISMDGSITVTLAAHQAIGLKIILAGTEE	
	OKAKYLPKLASGHIAAFCLTEPASGSDAASIRSRATLSDEKKHYILNGSKVWITNGG	
	LNIIFVFAKTEVYSDGSKDKITAFIVERDPGGVYNGKPEKDLGIRSNCEVHFE	
	NTKPIVENILGVDGDKVAMNINLSGRFSMGSVYAGLLKRLLEWTAETACRKOPNK	
	RLSEFGLIOEKFALMAOKAYVMESMTYLTAGMDLDQPFDPDCSEIAAMVYFSEAAWQ	
	CVSEALQILGGLGYTRDYPYERILRDRITRILLIPEGTNEILRMVYIALTGOHAGRILTT	
	RIHELQAKVSTVMDTVGRRLRDSLRGTDLGLTGNHVVHPSLADSANKFEENTYCF	
	GRVETLLLRFGKTIIMEOVLKRVANILINLYGMTAVLSRASRSIRIGLRNHDHVL	
	LANTFCVEALQNLFSLQSLDKVAPENLDQIKKVSQIILEKRAYICAHPLDRTC"	
BASE COUNT	471 a 481 c 519 g 392 t	
ORIGIN		
Query Match	100.0%; Score 1863; DB 6; Length 1863;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 1863; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 ATGAGCGGCTCGGGCTCTTCCTGCGCACACCGCTCGCGCTCGTGCCTCCCGGGTCTG 60	
Db	1 ATGAGCGGCTCGGGCTCTTCCTGCGCACACCGCTCGCGCTCGTGCCTCCCGGGTCTG 60	
Qy	61 GTGTGCTCTACCGGGAACCGGGCTACTCGGCACACCGCGCTGTACGAGCTTTTCGCC 120	
Db	61 GTGTGCTCTACCGGGAACCGGGCTACTCGGCACACCGCGCTGTACGAGCTTTTCGCC 120	
Qy	121 AAGAGCTTTTCTTAGGCAAAATCAAGAAGAAGTTCCTCCATTTCCAGAAAGTTAGC 180	
Db	121 AAGAGCTTTTCTTAGGCAAAATCAAGAAGAAGTTCCTCCATTTCCAGAAAGTTAGC 180	
Qy	181 CAAGATGACCTTAATGAATCAATCAGTCTCTGGGACCGCTGGAAAAATCTTCACATGAA 240	
Db	181 CAAGATGACCTTAATGAATCAATCAGTCTCTGGGACCGCTGGAAAAATCTTCACATGAA 240	
Qy	241 GAGTGGACTCCCGAAAAATTTGACAGGAGGAAAAATCCAGATGAATTTTGGAGAA 300	
Db	241 GAGTGGACTCCCGAAAAATTTGACAGGAGGAAAAATCCAGATGAATTTTGGAGAA 300	
Qy	301 TTGAAGACCTTAGGGCTTTTGGGCTCAAGTCCAGAGAAATATGCTGGCGTGGCTTC 360	
Db	301 TTGAAGACCTTAGGGCTTTTGGGCTCAAGTCCAGAGAAATATGCTGGCGTGGCTTC 360	
Qy	361 TCCAACCATGTACTCAAGACTAGGAGAGATCATCAGCATGGATGGGTCCATCACTGTG 420	
Db	361 TCCAACCATGTACTCAAGACTAGGAGAGATCATCAGCATGGATGGGTCCATCACTGTG 420	
Qy	421 ACCCTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCAGTGAAG 480	
Db	421 ACCCTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCAGTGAAG 480	
Qy	481 CAGAAAGCAATATTTCCTTAACTGGCGTCCGGGAGACATTTGACGCTTCTGGCTC 540	
Db	481 CAGAAAGCAATATTTCCTTAACTGGCGTCCGGGAGACATTTGACGCTTCTGGCTC 540	
Qy	541 ACGGAGCCAGCAGTGGAGCGATGCAGCTCAATTCGGAGCAGAGCCACACTAAGTGA 600	
Db	541 ACGGAGCCAGCAGTGGAGCGATGCAGCTCAATTCGGAGCAGAGCCACACTAAGTGA 600	
Qy	601 GACAGAGCAGCTACATCCTCAATGGCTCAAGCTCTGATTAATGAGGAGTGGCC 660	
Db	601 GACAGAGCAGCTACATCCTCAATGGCTCAAGCTCTGATTAATGAGGAGTGGCC 660	

Qy	661 AATATTTTACATGCTTTGCAAGAACTGAGGTCGTTGATTCTTGATGGATCAGTGAAGAC 720	
Db	661 AATATTTTACTGCTTTGCAAGAACTGAGGTCGTTGATTCTTGATGGATCAGTGAAGAC 720	
Qy	721 AAAATCACAGCATTCATAGTAGAAAAGAGACTTTGGTGAGTCACTAAATGGGAAACCCGAA 780	
Db	721 AAAATCACAGCATTCATAGTAGAAAAGAGACTTTGGTGAGTCACTAAATGGGAAACCCGAA 780	
Qy	781 GATAAATTAGGCATTCGGGGCTCCAACACTTGTGAAGTCCATTTTGAAGAACCAAGATA 840	
Db	781 GATAAATTAGGCATTCGGGGCTCCAACACTTGTGAAGTCCATTTTGAAGAACCAAGATA 840	
Qy	841 COTGTGAAAACATCCTTGCAGAGTTCGAGATGGTTTAAGGTGCCATGAACATCCTC 900	
Db	841 COTGTGAAAACATCCTTGCAGAGTTCGAGATGGTTTAAGGTGCCATGAACATCCTC 900	
Qy	901 AACAGCGCGGCTTCAGCATGGGACGCTGCTGGTGGCTGCTCAAGAGATTGATGAA 960	
Db	901 AACAGCGCGGCTTCAGCATGGGACGCTGCTGGTGGCTGCTCAAGAGATTGATGAA 960	
Qy	961 ATGACTGCTGAGTACGCCCTGCACAGGAACAGTTTAAACAGAGGCTCAGTCAATTGGA 1020	
Db	961 ATGACTGCTGAGTACGCCCTGCACAGGAACAGTTTAAACAGAGGCTCAGTCAATTGGA 1020	
Qy	1021 TTGATTCAGAGAAATTTGCACTGATGGCTCAGAAGCTTACGTATCGAGAGTATGACC 1080	
Db	1021 TTGATTCAGAGAAATTTGCACTGATGGCTCAGAAGCTTACGTATCGAGAGTATGACC 1080	
Qy	1081 TACCTCACAGCAGGATGCTGGACCAACTGGCTTTCCGAGCTGCTCAGAGGAGCC 1140	
Db	1081 TACCTCACAGCAGGATGCTGGACCAACTGGCTTTCCGAGCTGCTCAGAGGAGCC 1140	
Qy	1141 ATGGTGAAGTGTTCAGCTCCGAGGCGCTGCGAGTGTGAGTGAGGCGCTGACAGATC 1200	
Db	1141 ATGGTGAAGTGTTCAGCTCCGAGGCGCTGCGAGTGTGAGTGAGGCGCTGACAGATC 1200	
Qy	1201 CTCGGGGCTTGGGCTACACAAGGACTATCCGTACGAGCGCATCTGGGTGACACCCGC 1260	
Db	1201 CTCGGGGCTTGGGCTACACAAGGACTATCCGTACGAGCGCATCTGGGTGACACCCGC 1260	
Qy	1261 ATCCTCCTCATCTTCGAGGAAACCAATGAGATTCTCCGGATGTACATCGCCCTGACGGGT 1320	
Db	1261 ATCCTCCTCATCTTCGAGGAAACCAATGAGATTCTCCGGATGTACATCGCCCTGACGGGT 1320	
Qy	1321 CTCAGCATGCGCGCGCATCTTACTTACCAGGATCCATGAGCTTAAACAGGCGCAAGTG 1380	
Db	1321 CTCAGCATGCGCGCGCATCTTACTTACCAGGATCCATGAGCTTAAACAGGCGCAAGTG 1380	
Qy	1381 AGCACAGTCAATGATACCGTTGGCGGAGCTTTCGGGACTCCCTGGCGCACTGTGGAC 1440	
Db	1381 AGCACAGTCAATGATACCGTTGGCGGAGCTTTCGGGACTCCCTGGCGCACTGTGGAC 1440	
Qy	1441 CTGGGCTGACAGCAACCAATGAGTGTGACCCCACTTTCGGGAGCTGCGACAGTGCACAAG 1500	
Db	1441 CTGGGCTGACAGCAACCAATGAGTGTGACCCCACTTTCGGGAGCTGCGACAGTGCACAAG 1500	
Qy	1501 TTTGAGGAGAACACTTACTGCTTCGGCGGACCGTGGAGACACTGCTGCTCCGCTTTGGC 1560	
Db	1501 TTTGAGGAGAACACTTACTGCTTCGGCGGACCGTGGAGACACTGCTGCTCCGCTTTGGC 1560	
Qy	1561 AAGACCATCATGAGGAGCAGCTGCTGAAAGGGGTGGCCACACTCCTCATCAACCTG 1620	
Db	1561 AAGACCATCATGAGGAGCAGCTGCTGAAAGGGGTGGCCACACTCCTCATCAACCTG 1620	
Qy	1621 TATGGCATGAGCGCGTGTCTCGCGGCGCAGCGCTCCATCCGCAATTTGGGCTCCGCAAC 1680	
Db	1621 TATGGCATGAGCGCGTGTCTCGCGGCGCAGCGCTCCATCCGCAATTTGGGCTCCGCAAC 1680	
Qy	1681 CACGACCAAGGCTCTCTTTGGGCAACACCTTCTGCGTGGAAAGCTTACTTTCGAGAACTC 1740	
Db	1681 CACGACCAAGGCTCTCTTTGGGCAACACCTTCTGCGTGGAAAGCTTACTTTCGAGAACTC 1740	
Qy	1741 TTGACGCTCTCTCAGCTGGACAGTATGCTCCAGAAAACCTTAGATGAGCAGATTAAAGAA 1800	


```

|||||
1741 TTACGCTCTCTCAGCTGGACAAGTATGCTCCAGAAAACCTAGATGACAGATTAGAAA 1800
QY GTGTCCCGACGAGATCCTTTGAGACGAGCCTATATCTGTGCCACCTCTGGACAGACA 1860
Db GTGTCCCGACGAGATCCTTTGAGACGAGCCTATATCTGTGCCACCTCTGGACAGACA 1860
QY 1861 TGC 1863
Db 1861 TGC 1863

RESULT 2
BC013354
LOCUS
DEFINITION
  Homo sapiens very-long-chain acyl-CoA dehydrogenase VLCAD, mRNA
  (CDNA clone MGC:14970 IMAGE:3935925), complete cds.
ACCESSION
  BC013354
VERSION
  GI:15426497
KEYWORDS
  MGC.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Homo sapiens
  Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Eukaryota; Euteleostomi; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
  1 (bases 1 to 2448)
  Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
  Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
  Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
  Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Bhat,N.K.,
  Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
  Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
  Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshlyuk,S.,
  Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
  Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
  McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
  Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
  Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
  Sanchez,A., Whiting,M., Madan,A., Madan,A., Rodriguez,S.,
  Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
  Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
  Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E.,
  Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  Generation and initial analysis of more than 15,000 full-length
  human and mouse cDNA sequences
  Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  2338257
  12477932
  2 (bases 1 to 2448)
  Strausberg,R.
  Direct Submission
  Submitted (31-AUG-2001) National Institutes of Health, Mammalian
  Gene Collection (MGC), Cancer Genomics Office, National Cancer
  Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
  USA
  NIH-MGC Project URL: http://mgc.nci.nih.gov
  Contact: MGC help desk
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: DCTD/DTP
  CDNA Library Preparation: Rubin Laboratory
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Institute for Systems Biology
  http://www.systemsbio.org
  contact: amadan@systemsbiology.org
  Anup Madan, Jessica Fahey, Erin Helton, Mark Ketterman, Anuradha
  Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
  Clone distribution: MGC clone distribution information can be found
  through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
  Series: IRAL Plate: 25 Row: c Column: 21.
  Location/Qualifiers
    1..2448
    /organism="Homo sapiens"

```

```

/mol_type="mrna"
/db_xref="taxon:9606"
/clone="MGC:14970 IMAGE:3935925"
/lissue_type="Lung, small cell carcinoma"
/clone_lib="NIH_MGC_7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1..2448
/gene="VLCAD"
/note="synonyms: NP0002, MGC14452, ACAD-9"
/db_xref="LocusID:28976"
66..1931
/codon_start=1
/product="very-long-chain acyl-CoA dehydrogenase VLCAD"
/protein_id="AAH13354.1"
/db_xref="GI:15426498"
/db_xref="LocusID:28976"
/translacion="MSGGGLFRTTAARACRGLVWSTANRRLLRTSPVPRAPAKELF
LGKIKKKEVFPPEVSQDELNEINQFUGPVEKFFTEEVDSRKIDQEKIPDETLEKLR
SLGFLQVPEEYGGHIAAPCLTEPASGDAASIRSRATLSDEKHYILNGKIVTTNGE
QKAKYLPKLASGEHIAAPCLTEPASGDAASIRSRATLSDEKHYILNGKIVTTNGE
LANITFTFAKTEVVDSDGSKDKITAFIVERDGVGTNGKPEDKLGIRGNTCEVTFE
NKTIPVENILGEGVDGFKVAMNINLSGRFSMGSVVAGLLKRLIEMTAETACRQPNK
RLSEGLIQEKFALMAOKAVYMESMTYITAGLDQPGFPCDSIEAANYKVFSSAAWQ
CVSEALOILGGLGYTRDYPYERILDRITRILIFEGTNEILRMYIALTLQHQHAGILTT
RIHELKQKQSVTMDTVRRURDLSGRITVDLGTNGHGVHPSLADSAKFEENTYCF
LNTFTLLRFRGTIMBEQVLKRVANILINLYGMTAVLSRASRSIRIGLRNHDHVL
GATVEVAYLQNLFSQLDKYAPENLDEQIKKVSQIILEKRAYICAHPLDRTC"
BASE COUNT 624 a 617 c 682 g 525 t
ORIGIN
Query Match 100.0%; Score 1863; DB 9; Length 2448;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGCGGCTCGGGCTCTTCTCGGCACCAACGCGCTCGGCTCGTGGCTCGCGGGGCTG 60
Db 66 ATGAGCGGCTCGGGCTCTTCTCGGCACCAACGCGCTCGGCTCGTGGCTCGCGGGGCTG 125
QY 61 GTGCTCTTACCGCGAACCGCGGCTACTGGCACCAACGCGGCTGTACGAGCTTTGCGC 120
Db 126 GTGCTCTTACCGCGAACCGCGGCTACTGGCACCAACGCGGCTGTACGAGCTTTGCGC 185
QY 121 AAAGAGCTTTTCTTAGGCAAAATCAAGAAGAAAGAACTTTTCCATTTCAGAGTTCAG 180
Db 186 AAAGAGCTTTTCTTAGGCAAAATCAAGAAGAAAGAACTTTTCCATTTCAGAGTTCAG 245
QY 181 CAAGATGAATTAATGAAATCAATCAGTTCCTTGGGACCCCGTGGAAAAATTTCTTCACTGAA 240
Db 246 CAAGATGAATTAATGAAATCAATCAGTTCCTTGGGACCCCGTGGAAAAATTTCTTCACTGAA 305
QY 241 GAGTGGACCTCCGAAAAATTTGACGAGGAGGAAAAATCCAGATGAAACTTTGGAGAAA 300
Db 306 GAGTGGACCTCCGAAAAATTTGACGAGGAGGAAAAATCCAGATGAAACTTTGGAGAAA 365
QY 301 TTGAAGAGCTAGGGCTTTTGGGCTCAGAGTCCAGAGAAATATGTTGGCTCGGCTTC 360
Db 366 TTGAAGAGCTAGGGCTTTTGGGCTCAGAGTCCAGAGAAATATGTTGGCTCGGCTTC 425
QY 361 TCCAAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCATCTGTG 420
Db 426 TCCAAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCATCTGTG 485
QY 421 ACCCTGGCAGCGCACCGAGCTATTGGCTCAAGGGATCATCTTGGCTGGCACTGAGGAG 480
Db 486 ACCCTGGCAGCGCACCGAGCTATTGGCTCAAGGGATCATCTTGGCTGGCACTGAGGAG 545
QY 481 CAGAAAGCCAAATACTTGGCTAACTGGCTCCGGGAGACATATTGCAGCCTTCTGCCCTC 540
Db 546 CAGAAAGCCAAATACTTGGCTAACTGGCTCCGGGAGACATATTGCAGCCTTCTGCCCTC 605
QY 541 ACGGAGCCAGCAGTGGGAGCGCATCGAGCTCAATCCGGGAGAGCCACACTAAGTGA 600
|||||

```

Db	606	ACGGAGCCACCCAGTGGGAGCGATGAGCCTCAATCCGGAGCAGAGCCACACTAAGTGAA	665
Qy	601	GACAAGAAGCACTACATCCTCAATGGCTCCAAAGTCTCTGATTTACTTAATGGAGGACTGGCC	660
Db	666	GACAAGAGCACTACATCCTCAATGGCTCCAAAGTCTGATTTACTTAATGGAGGACTGGCC	725
Qy	661	AATATTTTTACTGTGTTTGCAAAGACTGAGGTCGTTGATTCTGATGGATCAGTGAAGAC	720
Db	726	AATATTTTTACTGTGTTTGCAAAGACTGAGGTCGTTGATTCTTCATGGATCAGTGAAGAC	785
Qy	721	AAATCACAGCACTCATAGTAGAAGAGACTTTGGTGGAGTCACATAATGGGAAACCCGAA	780
Db	786	AAATCACAGCACTCATAGTAGAAGAGACTTTGGTGGAGTCACATAATGGGAAACCCGAA	845
Qy	781	GATAAAATTAGGCATTCGGGCTCCCAACACTGTGTAAGTCCATTTTGAACACCAAGATA	840
Db	846	GATAAAATTAGGCATTCGGGCTCCCAACACTGTGTAAGTCCATTTTGAACACCAAGATA	905
Qy	841	CCTGTGAAACACATCCTTGGAGAGTCGGAGATGGGTTTAAAGTGGCCATGAACATCCTC	900
Db	906	CCTGTGAAACACATCCTTGGAGAGTCGGAGATGGGTTTAAAGTGGCCATGAACATCCTC	965
Qy	901	AACAGCGCCGGTTTCAGCATGGGCACGCGTGGCTGGCTGCTCAAGAGATTGATTGAA	960
Db	966	AACAGCGCCGGTTTCAGCATGGGCACGCGTGGCTGGCTGCTCAAGAGATTGATTGAA	1025
Qy	961	ATGACTGCTGAGTACGCCGTGCACAAGGAACAGTTTAAACAGAGGCTCAGTGAATTGGA	1020
Db	1026	ATGACTGCTGAGTACGCCGTGCACAAGGAACAGTTTAAACAGAGGCTCAGTGAATTGGA	1085
Qy	1021	TTGATTTCAGAGAAATTTGCACATGATGGGTCAGAAGGCTTACGTCATGAGAGTAGTAC	1080
Db	1086	TTGATTTCAGAGAAATTTGCACATGATGGGTCAGAAGGCTTACGTCATGAGAGTAGTAC	1145
Qy	1081	TACCTCACAGAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGCGAGC	1140
Db	1146	TACCTCACAGAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGCGAGC	1205
Qy	1141	ATGGTGAAGGTGTCAGCTCCGAGGCCGCTGGCAGTGTGTGATGAGGCGGTGCAAGATC	1200
Db	1206	ATGGTGAAGGTGTCAGCTCCGAGGCCGCTGGCAGTGTGTGATGAGGCGGTGCAAGATC	1265
Qy	1201	CTCGGGGCTTGGGCTACACAAGGGACTATCCGTACGAGGCGATCTGCGTGCACCCGC	1260
Db	1266	CTCGGGGCTTGGGCTACACAAGGGACTATCCGTACGAGGCGATCTGCGTGCACCCGC	1325
Qy	1261	ATCCTCTCATCTTCGAGGAAACCAATGAGATTCTCGGATGTACATCGCCCTGACGGGT	1320
Db	1326	ATCCTCTCATCTTCGAGGAAACCAATGAGATTCTCGGATGTACATCGCCCTGACGGGT	1385
Qy	1321	CTGCAGCATGCGGCGGATCCTGACTACAGAGATCCATGAGCTTAAACAGGCCAAGTG	1380
Db	1386	CTGCAGCATGCGGCGGATCCTGACTACAGAGATCCATGAGCTTAAACAGGCCAAGTG	1445
Qy	1381	AGCAGCATGAGATACGTTGGCGGAGGCTTCGGGACTCCCTGGGCCGAACTGTGGAC	1440
Db	1446	AGCAGCATGAGATACGTTGGCGGAGGCTTCGGGACTCCCTGGGCCGAACTGTGGAC	1505
Qy	1441	CTGGGGCTGACAGGCAACCATGGAGTTGTGCACCCCACTCTGGGACACGTGCCACAAG	1500
Db	1506	CTGGGGCTGACAGGCAACCATGGAGTTGTGCACCCCACTCTGGGACACGTGCCACAAG	1565
Qy	1501	TTTGAGGAGACACCTACTGCTTCGGCGGACCTGGAGACACTGCTGCTCCGCTTTGGC	1560
Db	1566	TTTGAGGAGACACCTACTGCTTCGGCGGACCTGGAGACACTGCTGCTCCGCTTTGGC	1625
Qy	1561	AAGACCATCATGGAGGACAGCTGCTACTGAAGCGGTTGGCCCAACATCCTCATCAAGCTG	1620
Db	1626	AAGACCATCATGGAGGACAGCTGCTACTGAAGCGGTTGGCCCAACATCCTCATCAAGCTG	1685
Qy	1621	TATGGCATGACGGCGTCTGTCCGGGCGACCGCTCCATCCGCATTTGGCTCCGCAAC	1680
Db	1686	TATGGCATGACGGCGTCTGTCCGGGCGACCGCTCCATCCGCATTTGGCTCCGCAAC	1745

Qy	1681	CACGACACAGAGTCTCTCTTGGCCACACACTTCTCGTGGAGAGCTTACTTGGCAGAAATCTC	1741
Db	1746	CACGACACAGAGTCTCTTGGCCACACACTTCTCGTGGAGAGCTTACTTGGCAGAAATCTC	1805
Qy	1741	TTGAGCTCTCTCAGCTGGAGCAAGTAGTGTCTCCAGAAAAACCTAGATGAGCAGATTAGAAA	1800
Db	1806	TTGAGCTCTCTCAGCTGGAGCAAGTAGTGTCTCCAGAAAAACCTAGATGAGCAGATTAGAAA	1865
Qy	1801	GTGTCCAGCAGATCTCTTGAGAGCGAGAGCTATATCTGTGCCACACCTCTGGACAGGACA	1860
Db	1866	GTGTCCAGCAGATCTCTTGAGAGCGAGAGCTATATCTGTGCCACACCTCTGGACAGGACA	1925
Qy	1861	TGC 1863	
Db	1926	TGC 1928	
RESULT 3	AX394747	Sequence 1 from Patent WO0218582.	
LOCUS	AX394747	2452 bp	DNA linear PAT 18-MAY-2002
DEFINITION	AX394747		
ACCESSION	AX394747		
VERSION	AX394747.1	GI:21065826	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			
Qy	1681	CACGACACAGAGTCTCTCTTGGCCACACACTTCTCGTGGAGAGCTTACTTGGCAGAAATCTC	1741
Db	1746	CACGACACAGAGTCTCTTGGCCACACACTTCTCGTGGAGAGCTTACTTGGCAGAAATCTC	1805
Qy	1741	TTGAGCTCTCTCAGCTGGAGCAAGTAGTGTCTCCAGAAAAACCTAGATGAGCAGATTAGAAA	1800
Db	1806	TTGAGCTCTCTCAGCTGGAGCAAGTAGTGTCTCCAGAAAAACCTAGATGAGCAGATTAGAAA	1865
Qy	1801	GTGTCCAGCAGATCTCTTGAGAGCGAGAGCTATATCTGTGCCACACCTCTGGACAGGACA	1860
Db	1866	GTGTCCAGCAGATCTCTTGAGAGCGAGAGCTATATCTGTGCCACACCTCTGGACAGGACA	1925
Qy	1861	TGC 1863	
Db	1926	TGC 1928	
RESULT 3	AX394747	Sequence 1 from Patent WO0218582.	
LOCUS	AX394747	2452 bp	DNA linear PAT 18-MAY-2002
DEFINITION	AX394747		
ACCESSION	AX394747		
VERSION	AX394747.1	GI:21065826	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			
Qy	1681	CACGACACAGAGTCTCTCTTGGCCACACACTTCTCGTGGAGAGCTTACTTGGCAGAAATCTC	1741
Db	1746	CACGACACAGAGTCTCTTGGCCACACACTTCTCGTGGAGAGCTTACTTGGCAGAAATCTC	1805
Qy	1741	TTGAGCTCTCTCAGCTGGAGCAAGTAGTGTCTCCAGAAAAACCTAGATGAGCAGATTAGAAA	1800
Db	1806	TTGAGCTCTCTCAGCTGGAGCAAGTAGTGTCTCCAGAAAAACCTAGATGAGCAGATTAGAAA	1865
Qy	1801	GTGTCCAGCAGATCTCTTGAGAGCGAGAGCTATATCTGTGCCACACCTCTGGACAGGACA	1860
Db	1866	GTGTCCAGCAGATCTCTTGAGAGCGAGAGCTATATCTGTGCCACACCTCTGGACAGGACA	1925
Qy	1861	TGC 1863	
Db	1926	TGC 1928	
RESULT 3	AX394747	Sequence 1 from Patent WO0218582.	
LOCUS	AX394747	2452 bp	DNA linear PAT 18-MAY-2002
DEFINITION	AX394747		
ACCESSION	AX394747		
VERSION	AX394747.1	GI:21065826	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			
Qy	1681	CACGACACAGAGTCTCTCTTGGCCACACACTTCTCGTGGAGAGCTTACTTGGCAGAAATCTC	1741
Db	1746	CACGACACAGAGTCTCTTGGCCACACACTTCTCGTGGAGAGCTTACTTGGCAGAAATCTC	1805
Qy	1741	TTGAGCTCTCTCAGCTGGAGCAAGTAGTGTCTCCAGAAAAACCTAGATGAGCAGATTAGAAA	1800
Db	1806	TTGAGCTCTCTCAGCTGGAGCAAGTAGTGTCTCCAGAAAAACCTAGATGAGCAGATTAGAAA	1865
Qy	1801	GTGTCCAGCAGATCTCTTGAGAGCGAGAGCTATATCTGTGCCACACCTCTGGACAGGACA	1860
Db	1866	GTGTCCAGCAGATCTCTTGAGAGCGAGAGCTATATCTGTGCCACACCTCTGGACAGGACA	1925
Qy	1861	TGC 1863	
Db	1926	TGC 1928	
RESULT 3	AX394747	Sequence 1 from Patent WO0218582.	
LOCUS	AX394747	2452 bp	DNA linear PAT 18-MAY-2002
DEFINITION	AX394747		
ACCESSION	AX394747		
VERSION	AX394747.1	GI:21065826	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			
Qy	1681	CACGACACAGAGTCTCTCTTGGCCACACACTTCTCGTGGAGAGCTTACTTGGCAGAAATCTC	1741
Db	1746	CACGACACAGAGTCTCTTGGCCACACACTTCTCGTGGAGAGCTTACTTGGCAGAAATCTC	1805
Qy	1741	TTGAGCTCTCTCAGCTGGAGCAAGTAGTGTCTCCAGAAAAACCTAGATGAGCAGATTAGAAA	1800
Db	1806	TTGAGCTCTCTCAGCTGGAGCAAGTAGTGTCTCCAGAAAAACCTAGATGAGCAGATTAGAAA	1865
Qy	1801	GTGTCCAGCAGATCTCTTGAGAGCGAGAGCTATATCTGTGCCACACCTCTGGACAGGACA	1860
Db</			

181 CAAGATGAACCTTAATGAAATCAATCAGTTCTTGGACCCGTGAAAAATTCCTCACTGAA 240
182 |
183 |
184 |
185 |
186 |
187 |
188 |
189 |
190 |
191 |
192 |
193 |
194 |
195 |
196 |
197 |
198 |
199 |
200 |
201 |
202 |
203 |
204 |
205 |
206 |
207 |
208 |
209 |
210 |
211 |
212 |
213 |
214 |
215 |
216 |
217 |
218 |
219 |
220 |
221 |
222 |
223 |
224 |
225 |
226 |
227 |
228 |
229 |
230 |
231 |
232 |
233 |
234 |
235 |
236 |
237 |
238 |
239 |
240 |
241 GAGGTGACTCCCGAAAAATTTGACAGGAGGAAAAATCCAGATGAAATTTGGAGAAA 300
242 |
243 |
244 |
245 |
246 |
247 CAAGATGAACCTTAATGAAATCAATCAGTTCTTGGACCCGTGAAAAATTCCTCACTGAA 306
248 |
249 |
250 |
251 |
252 |
253 |
254 |
255 |
256 |
257 |
258 |
259 |
260 |
261 |
262 |
263 |
264 |
265 |
266 |
267 |
268 |
269 |
270 |
271 |
272 |
273 |
274 |
275 |
276 |
277 |
278 |
279 |
280 |
281 |
282 |
283 |
284 |
285 |
286 |
287 |
288 |
289 |
290 |
291 |
292 |
293 |
294 |
295 |
296 |
297 |
298 |
299 |
300 |
301 TTCAAGAGCCTACGGCTTTTGGGCTGCAAGTCCCAAGAGATATGTTGGCTTGGCTTC 360
302 |
303 |
304 |
305 |
306 |
307 GAGGTGACTCCCGAAAAATTTGACAGGAGGAAAAATCCAGATGAAATTTGGAGAAA 366
308 |
309 |
310 |
311 |
312 |
313 |
314 |
315 |
316 |
317 |
318 |
319 |
320 |
321 |
322 |
323 |
324 |
325 |
326 |
327 |
328 |
329 |
330 |
331 |
332 |
333 |
334 |
335 |
336 |
337 |
338 |
339 |
340 |
341 |
342 |
343 |
344 |
345 |
346 |
347 |
348 |
349 |
350 |
351 |
352 |
353 |
354 |
355 |
356 |
357 |
358 |
359 |
360 |
361 TCCAACACCATGTAACAGACTAGGGAGATCATCAGCATGATGATGGTCCATCACTGTG 420
362 |
363 |
364 |
365 |
366 |
367 TTCAAGAGCCTACGGCTTTTGGGCTGCAAGTCCCAAGAGATATGTTGGCTTGGCTTC 426
368 |
369 |
370 |
371 |
372 |
373 |
374 |
375 |
376 |
377 |
378 |
379 |
380 |
381 |
382 |
383 |
384 |
385 |
386 |
387 |
388 |
389 |
390 |
391 |
392 |
393 |
394 |
395 |
396 |
397 |
398 |
399 |
400 |
401 |
402 |
403 |
404 |
405 |
406 |
407 |
408 |
409 |
410 |
411 |
412 |
413 |
414 |
415 |
416 |
417 |
418 |
419 |
420 |
421 ACCCTGGCAGCCAGCAGGCTATTTGGCTCAAGGGGATCATCTTGGCTGGCCTGAGGAG 480
422 |
423 |
424 |
425 |
426 |
427 TCCAACACCATGTAACAGACTAGGGAGATCATCAGCATGATGATGGTCCATCACTGTG 486
428 |
429 |
430 |
431 |
432 |
433 |
434 |
435 |
436 |
437 |
438 |
439 |
440 |
441 |
442 |
443 |
444 |
445 |
446 |
447 |
448 |
449 |
450 |
451 |
452 |
453 |
454 |
455 |
456 |
457 |
458 |
459 |
460 |
461 |
462 |
463 |
464 |
465 |
466 |
467 |
468 |
469 |
470 |
471 |
472 |
473 |
474 |
475 |
476 |
477 |
478 |
479 |
480 |
481 CAGAAAGCCAAATACCTTAACTGGGCTCCGGGAGCAGATTCAGTCCGAGGAGGAGTGA 600
482 |
483 |
484 |
485 |
486 |
487 ACCCTGGCAGCCAGCAGGCTATTTGGCTCAAGGGGATCATCTTGGCTGGCCTGAGGAG 546
488 |
489 |
490 |
491 |
492 |
493 |
494 |
495 |
496 |
497 |
498 |
499 |
500 |
501 |
502 |
503 |
504 |
505 |
506 |
507 |
508 |
509 |
510 |
511 |
512 |
513 |
514 |
515 |
516 |
517 |
518 |
519 |
520 |
521 |
522 |
523 |
524 |
525 |
526 |
527 |
528 |
529 |
530 |
531 |
532 |
533 |
534 |
535 |
536 |
537 |
538 |
539 |
540 |
541 ACGGAGCCAGCAGTGGGAGGATGAGCCTCAATCCGAGGAGGAGGAGTGA 600
542 |
543 |
544 |
545 |
546 |
547 CAGAAAGCCAAATACCTTAACTGGGCTCCGGGAGCAGATTCAGTCCGAGGAGGAGTGA 606
548 |
549 |
550 |
551 |
552 |
553 |
554 |
555 |
556 |
557 |
558 |
559 |
560 |
561 |
562 |
563 |
564 |
565 |
566 |
567 |
568 |
569 |
570 |
571 |
572 |
573 |
574 |
575 |
576 |
577 |
578 |
579 |
580 |
581 |
582 |
583 |
584 |
585 |
586 |
587 |
588 |
589 |
590 |
591 |
592 |
593 |
594 |
595 |
596 |
597 |
598 |
599 |
600 |
601 GACAAGACACTACATCTCAATGGCTCCAGGCTGAGTAACTAGTAACTAGTAACTAGTAA 660
602 |
603 |
604 |
605 |
606 |
607 ACGGAGCCAGCAGTGGGAGGATGAGCCTCAATCCGAGGAGGAGGAGTGA 666
608 |
609 |
610 |
611 |
612 |
613 |
614 |
615 |
616 |
617 |
618 |
619 |
620 |
621 |
622 |
623 |
624 |
625 |
626 |
627 |
628 |
629 |
630 |
631 |
632 |
633 |
634 |
635 |
636 |
637 |
638 |
639 |
640 |
641 |
642 |
643 |
644 |
645 |
646 |
647 |
648 |
649 |
650 |
651 AATATTTTACTGTGTTGCAAGACTGAGGTGCTTGAATTCATGATGAGTCACTGAAAGAC 720
652 |
653 |
654 |
655 |
656 |
657 |
658 |
659 |
660 |
661 |
662 |
663 |
664 |
665 |
666 |
667 |
668 |
669 |
670 |
671 |
672 |
673 |
674 |
675 |
676 |
677 |
678 |
679 |
680 |
681 |
682 |
683 |
684 |
685 |
686 |
687 |
688 |
689 |
690 |
691 |
692 |
693 |
694 |
695 |
696 |
697 |
698 |
699 |
700 |
701 |
702 |
703 |
704 |
705 |
706 |
707 |
708 |
709 |
710 |
711 |
712 |
713 |
714 |
715 |
716 |
717 |
718 |
719 |
720 |
721 AAATTCACAGCATTCATAGTAGAAAGACATTTGGTGGAGTCACTAAATGGGAAACCGAA 780
722 |
723 |
724 |
725 |
726 |
727 AATATTTTACTGTGTTGCAAGACTGAGGTGCTTGAATTCATGATGAGTCACTGAAAGAC 786
728 |
729 |
730 |
731 |
732 |
733 |
734 |
735 |
736 |
737 |
738 |
739 |
740 |
741 |
742 |
743 |
744 |
745 |
746 |
747 |
748 |
749 |
750 |
751 |
752 |
753 |
754 |
755 |
756 |
757 |
758 |
759 |
760 |
761 |
762 |
763 |
764 |
765 |
766 |
767 |
768 |
769 |
770 |
771 |
772 |
773 |
774 |
775 |
776 |
777 |
778 |
779 |
780 |
781 GATAAATAGGCATTCGGGCTCCCAACACTTCTGAAGTCCATTTTGAACACCAAGATA 840
782 |
783 |
784 |
785 |
786 |
787 AAATTCACAGCATTCATAGTAGAAAGACATTTGGTGGAGTCACTAAATGGGAAACCGAA 846
788 |
789 |
790 |
791 |
792 |
793 |
794 |
795 |
796 |
797 |
798 |
799 |
800 |
801 |
802 |
803 |
804 |
805 |
806 |
807 |
808 |
809 |
810 |
811 |
812 |
813 |
814 |
815 |
816 |
817 |
818 |
819 |
820 |
821 |
822 |
823 |
824 |
825 |
826 |
827 |
828 |
829 |
830 |
831 |
832 |
833 |
834 |
835 |
836 |
837 |
838 |
839 |
840 |
841 CCTGTGAAACATCCTTGGAGAGCTCGGAGATGGGTTTAAAGTGGCCATGAACATCTC 900
842 |
843 |
844 |
845 |
846 |
847 GATAAATAGGCATTCGGGCTCCCAACACTTGTGAAGTCCATTTTGAACACCAAGATA 906
848 |
849 |
850 |
851 |
852 |
853 |
854 |
855 |
856 |
857 |
858 |
859 |
860 |
861 |
862 |
863 |
864 |
865 |
866 |
867 |
868 |
869 |
870 |
871 |
872 |
873 |
874 |
875 |
876 |
877 |
878 |
879 |
880 |
881 |
882 |
883 |
884 |
885 |
886 |
887 |
888 |
889 |
890 |
891 |
892 |
893 |
894 |
895 |
896 |
897 |
898 |
899 |
900 |
901 AACACGCGCGGTTACAGATGGCAGCGCTGCTGGCTGGCTGCTCAAGAGATTTGATTGAA 960
902 |
903 |
904 |
905 |
906 |
907 CCTGTGAAACATCCTTGGAGAGCTCGGAGATGGGTTTAAAGTGGCCATGAACATCTC 966
908 |
909 |
910 |
911 |
912 |
913 |
914 |
915 |
916 |
917 |
918 |
919 |
920 |
921 |
922 |
923 |
924 |
925 |
926 |
927 |
928 |
929 |
930 |
931 |
932 |
933 |
934 |
935 |
936 |
937 |
938 |
939 |
940 |
941 |
942 |
943 |
944 |
945 |
946 |
947 |
948 |
949 |
950 |
951 |
952 |
953 |
954 |
955 |
956 |
957 |
958 |
959 |
960 |
961 ATGACTGTGAGTACGCTGCAAGAGGAAACAGTTTAAAGAGGCTCAGTGAATTTGGA 1020
962 |
963 |
964 |
965 |
966 |
967 AACACGCGCGGTTACAGATGGCAGCGCTGCTGGCTGGCTGCTCAAGAGATTTGATTGAA 1026
968 |
969 |
970 |
971 |
972 |
973 |
974 |
975 |
976 |
977 |
978 |
979 |
980 |
981 |
982 |
983 |
984 |
985 |
986 |
987 |
988 |
989 |
990 |
991 |
992 |
993 |
994 |
995 |
996 |
997 |
998 |
999 |
1000 |
1001 |
1002 |
1003 |
1004 |
1005 |
1006 |
1007 |
1008 |
1009 |
1010 |
1011 |
1012 |
1013 |
1014 |
1015 |
1016 |
1017 |
1018 |
1019 |
1020 |
1021 TTGATTCAGGAGAAATTTGCACTGATGCTCAGAGGCTTACGTCATGGAGATGATGACC 1080
1022 |
1023 |
1024 |
1025 |
1026 |
1027 ATGACTGTGAGTACGCTGCAAGAGGAAACAGTTTAAAGAGGCTCAGTGAATTTGGA 1086
1028 |
1029 |
1030 |
1031 |
1032 |
1033 |
1034 |
1035 |
1036 |
1037 |
1038 |
1039 |
1040 |
1041 |
1042 |
1043 |
1044 |
1045 |
1046 |
1047 |
1048 |
1049 |
1050 |
1051 |
1052 |
1053 |
1054 |
1055 |
1056 |
1057 |
1058 |
1059 |
1060 |
1061 |
1062 |
1063 |
1064 |
1065 |
1066 |
1067 |
1068 |
1069 |
1070 |
1071 |
1072 |
1073 |
1074 |
1075 |
1076 |
1077 |
1078 |
1079 |
1080 |
1081 TACCTCAGCAGGAGATGCTGACCAACCTTCCGACTGCTCCATCGAGGAGCC 1140
1082 |
1083 |
1084 |
1085 |
1086 |
1087 TTGATTCAGGAGAAATTTGCACTGATGCTCAGAGGCTTACGTCATGGAGATGATGACC 1146
1088 |
1089 |
1090 |
1091 |
1092 |
1093 |
1094 |
1095 |
1096 |
1097 |
1098 |
1099 |
1100 |
1101 |
1102 |
1103 |
1104 |
1105 |
1106 |
1107 |
1108 |
1109 |
1110 |
1111 |
1112 |
1113 |
1114 |
1115 |
1116 |
1117 |
1118 |
1119 |
1120 |
1121 |
1122 |
1123 |
1124 |
1125 |
1126 |
1127 |
1128 |
1129 |
1130 |
1131 |
1132 |
1133 |
1134 |
1135 |
1136 |
1137 |
1138 |
1139 |
1140 |
1141 ATGTTGAAGTGTTCAGCTCCGAGCGGCTTGGCAGTGTGTGAGTGAAGGCTGCAGATC 1200
1142 |
1143 |
1144 |
1145 |
1146 |
1147 TACCTCAGCAGGAGATGCTGGACCACTTCCGACTGCTCCATCGAGGAGCC 1206
1148 |
1149 |
1150 |
1151 |
1152 |
1153 |
1154 |
1155 |
1156 |
1157 |
1158 |
1159 |
1160 |
1161 |
1162 |
1163 |
1164 |
1165 |
1166 |
1167 |
1168 |
1169 |
1170 |
1171 |
1172 |
1173 |
1174 |
1175 |
1176 |
1177 |
1178 |
1179 |
1180 |
1181 |
1182 |
1183 |
1184 |
1185 |
1186 |
1187 |
1188 |
1189 |
1190 |
1191 |
1192 |
1193 |
1194 |
1195 |
1196 |
1197 |
1198 |
1199 |
1200 |
1201 CTCGGGGCTTGGCTACACAGGAGTATCCGTACGAGGCGATACCTGCGTGACACCCGC 1260
1202 |
1203 |
1204 |
1205 |
1206 |
1207 ATGTTGAAGTGTTCAGCTCCGAGCGGCTTGGCAGTGTGTGAGTGAAGGCTGCAGATC 1266
1208 |
1209 |
1210 |
1211 |
1212 |
1213 |
1214 |
1215 |
1216 |
1217 |
1218 |
1219 |
1220 |
1221 |
1222 |
1223 |
1224 |
1225 |
1226 |
1227 |
1228 |
1229 |
1230 |
1231 |
1232 |
1233 |
1234 |
1235 |
1236 |
1237 |
1238 |
1239 |
1240 |
1241 |
1242 |
1243 |
1244 |
1245 |
1246 |
1247 |
1248 |
1249 |
1250 |
1251 |
1252 |
1253 |
1254 |
1255 |
1256 |
1257 |
1258 |
1259 |
1260 |
1261 |
1262 |
1263 |
1264 |
1265 |
1266 |
1267 CTCGGGGCTTGGCTACACAGGAGTATCCGTACGAGGCGATACCTGCGTGACACCCGC 1326
1268 |
1269 |
1270 |
1271 |
1272 |
1273 |
1274 |
1275 |
1276 |
1277 |
1278 |
1279 |
1280 |
1281 |
1282 |
1283 |
1284 |
1285 |
1286 |
1287 |
1288 |
1289 |
1290 |
1291 |
1292 |
1293 |
1294 |
1295 |
1296 |
1297 |
1298 |
1299 |
1300 |
1301 |
1302 |
1303 |
1304 |
1305 |
1306 |
1307 |
1308 |
1309 |
1310 |
1311 |
1312 |
1313 |
1314 |
1315 |
1316 |
1317 |
1318 |
1319 |
1320 |
1321 |
1322 |
1323 |
1324 |
1325 |
1326 |
1327 |
1328 |
1329 |
1330 |
1331 |
1332 |
1333 |
1334 |
1335 |
1336 |
1337 |
1338 |
1339 |
1340 |
1341 |
1342 |
1343 |
1344 |
1345 |
1346 |
1347 |
1348 |
1349 |
1350 |
1351 |
1352 |
1353 |
1354 |
1355 |
1356 |
1357 |
1358 |
1359 |
1360 |
1361 |
1362 |
1363 |
1364 |
1365 |
1366 |
1367 |
1368 |
1369 |
1370 |
1371 |
1372 |
1373 |
1374 |
1375 |
1376 |
1377 |
1378 |
1379 |
1380 |
1381 |
1382 |
1383 |
1384 |
1385 |
1386 |
1387 |
1388 |
1389 |
1390 |
1391 |
1392 |
1393 |
1394 |
1395 |
1396 |
1397 |
1398 |
1399 |
1400 |
1401 |
1402 |
1403 |
1404 |
1405 |
1406 |
1407 |
1408 |
1409 |
1410 |
1411 |
1412 |
1413 |
1414 |
1415 |
1416 |
1417 |
1418 |
1419 |
1420 |
1421 |
1422 |
1423 |
1424 |
1425 |
1426 |
1427 |
1428 |
1429 |
1430 |
1431 |
1432 |
1433 |
1434 |
1435 |
1436 |
1437 |
1438 |
1439 |
1440 |
1441 |
1442 |
1443 |
1444 |
1445 |
1446 |
1447 |
1448 |
1449 |
1450 |
1451 |
1452 |
1453 |
1454 |
1455 |
1456 |
1457 |
1458 |
1459 |
1460 |
1461 |
1462 |
1463 |
1464 |
1465 |
1466 |
1467 |
1468 |
1469 |
1470 |
1471 |
1472 |
1473 |
1474 |
1475 |
1476 |
1477 |
1478 |
1479 |
1480 |
1481 |
1482 |
1483 |
1484 |
1485 |
1486 |
1487 |
1488 |
1489 |
1490 |
1491 |
1492 |
1493 |
1494 |
1495 |
1496 |
1497 |
1498 |
1499 |
1500 |
1501 |
1502 |
1503 |
1504 |
1505 |
1506 |
1507 |
1508 |
1509 |
1510 |
1511 |
1512 |
1513 |
1514 |
1515 |
1516 |
1517 |
1518 |
1519 |
1520 |
1521 |
1522 |
1523 |
1524 |
1525 |
1526 |
1527 |
1528 |
1529 |
1530 |
1531 |
1532 |
1533 |
1534 |
1535 |
1536 |
1537 |
1538 |
1539 |
1540 |
1541 |
1542 |
1543 |
1544 |
1545 |
1546 |
1547 |
1548 |
1549 |
1550 |
1551 |
1552 |
1553 |
1554 |
1555 |
1556 |
1557 |
1558 |
1559 |
1560 |
1561 |
1562 |
1563 |
1564 |
1565 |
1566 |
1567 |
1568 |
1569 |
1570 |
1571 |
1572 |
1573 |
1574 |
1575 |
1576 |
1577 |
1578 |
1579 |
1580 |
1581 |
1582 |
1583 |
1584 |
1585 |
1586 |
1587 |
1588 |
1589 |
1590 |
1591 |
1592 |
1593 |
1594 |
1595 |
1596 |
1597 |
1598 |
1599 |
1600 |
1601 |
1602 |
1603 |
1604 |
1605 |
1606 |
1607 |
1608 |
1609 |
1610 |
1611 |
1612 |
1613 |
1614 |
1615 |
1616 |
1617 |
1618 |
1619 |
1620 |
1621 |
1622 |
1623 |
1624 |
1625 |
1626 |
1627 |
1628 |
1629 |
1630 |
1631 |
1632 |
1633 |
1634 |
1635 |
1636 |
1637 |
1638 |
1639 |
1640 |
1641 |
1642 |
1643 |
1644 |
1645 |
1646 |
1647 |
1648 |
1649 |
1650 |
1651 |
1652 |
1653 |
1654 |
1655 |
1656 |
1657 |
1658 |
1659 |
1660 |
1661 |
1662 |
1663 |
1

PC	C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00	CC
Primer for synthesizing full-length cDNA and use thereof FH Key		
FT	CDS	Location/Qualifiers
FEATURES	source	1..2399
		/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"
BASE COUNT	597 a	606 c 672 g 524 t
ORIGIN		
Query Match 99.8%; Score 1859.8; DB 6; Length 2399;		
Best Local Similarity 99.9%; Pred. No. 0;		
Matches 1861; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
QY	1	ATGAGCGGCTGGGGCTCTTCTTGGCCACCAAGGCTGCGGCTCGTGCCTGCGGGGTCTG 60
DB	47	ATGAGCGGCTGGGGCTCTTCTTGGCCACCAAGGCTGCGGCTCGTGCCTGCGGGGTCTG 106
QY	61	GTGGTCTTACCGCGAACCGGGCTTACTGCGCACCAGCCAGCCCGCTGTAGAGCTTTCGCC 120
DB	107	GTGGTCTTACCGCGAACCGGGCTTACTGCGCACCAGCCAGCCCGCTGTAGAGCTTTCGCC 166
QY	121	AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGTTTTCCCAATTTCCAGAAAGTTAGC 180
DB	167	AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGTTTTCCCAATTTCCAGAAAGTTAGC 226
QY	181	CAAGATGAACCTTAATGAATCAATAGTTCCTTGGGACCCGCTGGAAAAATTTCTTCACTGAA 240
DB	227	CAAGATGAACCTTAATGAATCAATAGTTCCTTGGGACCCGCTGGAAAAATTTCTTCACTGAA 286
QY	241	GAGTGGACTCCGAAAAATTGACAGGAGGAGGAAAAATCCAGATGAAATTTTGGAGAAA 300
DB	287	GAGTGGACTCCGAAAAATTGACAGGAGGAGGAAAAATCCAGATGAAATTTTGGAGAAA 346
QY	301	TTCAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCCAGAAAGTAATGGTGGCCTGGGCTTC 360
DB	347	TTCAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCCAGAAAGTAATGGTGGCCTGGGCTTC 406
QY	361	TCCAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCACTGTG 420
DB	407	TCCAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCACTGTG 466
QY	421	ACCTTGCAGCGCACCAGGCTATTGGGCTCAAGGGGATCATCTTGGCTGGCACTGAGGAG 480
DB	467	ACCTTGCAGCGCACCAGGCTATTGGGCTCAAGGGGATCATCTTGGCTGGCACTGAGGAG 526
QY	481	CAGAAAGCCAAATACCTTGCCTAACTGGCGTCCGGGAGCACATTCAGGCTTCTTGCTC 540
DB	527	CAGAAAGCCAAATACCTTGCCTAACTGGCGTCCGGGAGCACATTCAGGCTTCTTGCTC 586
QY	541	ACGGAGCCAGCAGTGGAGCGATGACGCTCAATCCGGAGCAGAGCCACACTAAAGTAA 600
DB	587	ACGGAGCCAGCAGTGGAGCGATGACGCTCAATCCGGAGCAGAGCCACACTAAAGTAA 646
QY	601	GACAAGAGCACTACATCTCAATGGCTCCAAAGTCTGGATTTACTAATGAGGAGCTGGCC 660
DB	647	GACAAGAGCACTACATCTCAATGGCTCCAAAGTCTGGATTTACTAATGAGGAGCTGGCC 706
QY	661	AATATTTTACTGTGTTTGAAGAGCTGAGGTCTGATTTCTGATGGATCAGTGAAGAGAC 720
DB	707	AATATTTTACTGTGTTTGAAGAGCTGAGGTCTGATTTCTGATGGATCAGTGAAGAGAC 766
QY	721	AAATATCAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAA 780
DB	767	AAATATCAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAA 826
QY	781	GATAAAATTAGGCATTCGGGGCTCCAAACACTTGTGAAGTCCATTTTGAACACCAAGATA 840
DB	827	GATAAAATTAGGCATTCGGGGCTCCAAACACTTGTGAAGTCCATTTTGAACACCAAGATA 886
QY	841	CCTGTGAAAAACATCTTGGAGAGGTGGAGATGGGTTTAAAGTGGCCATGAACATCCTC 900

DB	887	CCTGTGAAAAACATCCTTTGGAGAGGTCCGAGATGGTTTAAAGTGGCCATGAACATCCTC 946
QY	901	AACAGCGCCGGTTTCAGCATGGCAGCGTCTGCTGGCTGCTCAAGAGATTGATTGAA 960
DB	947	AACAGCGCCGGTTTCAGCATGGCAGCGTCTGCTGGCTGCTCAAGAGATTGATTGAA 1006
QY	961	ATGACTGCTGAGTACGCTGCACAGAGAAACAGTTTAAACAAGAGGCTCAGTGAATTTGGA 1020
DB	1007	ATGACTGCTGAGTACGCTGCACAGAGAAACAGTTTAAACAAGAGGCTCAGTGAATTTGGA 1066
QY	1021	TTGATTCAGAGAAATTTGCACTGATGGCTCAGAGGCTTACGTCATCGAGAGTATGACC 1080
DB	1067	TTGATTCAGAGAAATTTGCACTGATGGCTCAGAGGCTTACGTCATCGAGAGTATGACC 1126
QY	1081	TACCTTCACAGCAGGATGCTGGACCAAGCTGCTTCCGACTGCTCCATCGAGCAGCC 1140
DB	1127	TACCTTCACAGCAGGATGCTGGACCAAGCTGCTTCCGACTGCTCCATCGAGCAGCC 1186
QY	1141	ATGSGTGAAGTGTTCAGCTCCGAGCCGCTGGCAGTGTGTGAGTGAGCGCTGCAGATC 1200
DB	1187	ATGSGTGAAGTGTTCAGCTCCGAGCCGCTGGCAGTGTGTGAGTGAGCGCTGCAGATC 1246
QY	1201	CTCGGGGCTTGGGCTACACAGGAGCTATCCGTACGAGCGCATCTGCTGACACCCGC 1260
DB	1247	CTCGGGGCTTGGGCTACACAGGAGCTATCCGTACGAGCGCATCTGCTGACACCCGC 1306
QY	1261	ATCCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGATGTACATCCCTGACGGGT 1320
DB	1307	ATCCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGATGTACATCCCTGACGGGT 1366
QY	1321	CTGAGCATGCGCGGCTATCTGACTACAGGATCCATGAGCTTAAACAGGCCAAAGTG 1380
DB	1367	CTGAGCATGCGCGGCTATCTGACTACAGGATCCATGAGCTTAAACAGGCCAAAGTG 1426
QY	1381	AGCAGCATCATGATACCGTTGGCGGAGCTTCCGAGCTCCCTGGCGGAACTGTGGAC 1440
DB	1427	AGCAGCATCATGATACCGTTGGCGGAGCTTCCGAGCTCCCTGGCGGAACTGTGGAC 1486
QY	1441	CTGGGCTCAGAGGCAACCATGGAGTTGTGACCCCGAGTCTTCCGAGCATGCCACAAG 1500
DB	1487	CTGGGCTCAGAGGCAACCATGGAGTTGTGACCCCGAGTCTTCCGAGCATGCCACAAG 1546
QY	1501	TTTGAGGAGAACACCTACTGCTTCCGCGGACCTGGAGACACTGCTGCTCCGCTTTGGC 1560
DB	1547	TTTGAGGAGAACACCTACTGCTTCCGCGGACCTGGAGACACTGCTGCTCCGCTTTGGC 1606
QY	1561	AAGACCATCATGGAGGAGCAGTGTGACTGAAGGGGTGGCCAACTCTCATCAACCTG 1620
DB	1607	AAGACCATCATGGAGGAGCAGTGTGACTGAAGGGGTGGCCAACTCTCATCAACCTG 1666
QY	1621	TATGGCATGACGCCCTGCTGCTCGCGGCGCAGCTCCATCCGCTTGGGCTCCGCAAC 1680
DB	1667	TATGGCATGACGCCCTGCTGCTCGCGGCGCAGCTCCATCCGCTTGGGCTCCGCAAC 1726
QY	1681	CAGCAGCAGAGTTCTCTTTGGCCAACTCTTCTGCTGCTGAGAGCTTACTTTGAGAAATCTC 1740
DB	1727	CAGCAGCAGAGTTCTCTTTGGCCAACTCTTCTGCTGCTGAGAGCTTACTTTGAGAAATCTC 1786
QY	1741	TTGAGGCTCTCTCAGCTGGCAAGTATGCTCCAGAAACCTAGATGAGCAGATTAGAAA 1800
DB	1787	TTGAGGCTCTCTCAGCTGGCAAGTATGCTCCAGAAACCTAGATGAGCAGATTAGAAA 1846
QY	1801	GTCTCCAGCAGATCTTGAAGAAGCAGCTATATCTGTGCCACCTCTTGGACAGGACA 1860
DB	1847	GTCTCCAGCAGATCTTGAAGAAGCAGCTATATCTGTGCCACCTCTTGGACAGGACA 1906
QY	1861	TGC 1863
DB	1907	TGC 1909

AK024012	LOCUS	AK024012	2399 bp	mRNA	linear	PRI 01-AUG-2002
DEFINITION	Homo sapiens cDNA FLJ13950 fis, clone Y79AA1001048, weakly similar to ACYL-COA DEHYDROGENASE, VERY-LONG-CHAIN SPECIFIC PRECURSOR (EC 1.3.99.-).					
ACCESSION	AK024012					
VERSION	AK024012.1 GI:10436257					
KEYWORDS	oligo capping; fis (full insert sequence).					
SOURCE	Homo sapiens					
ORGANISM	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wagatsuma, M., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T. NEDO human cDNA sequencing project					
AUTHORS	Unpublished					
REFERENCE	2 (bases 1 to 2399)					
TITLE	Isogai, T. and Otsuki, T.					
JOURNAL	Direct Submission					
AUTHORS	Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)					
JOURNAL	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.					
COMMENT	Location/Qualifiers					
FEATURES	1..2399					
source	/organism="Homo sapiens"					
	/mol_type="mRNA"					
	/db_xref="taxon:9606"					
	/clone="Y79AA1001048"					
	/cell_line="Y79"					
	/cell_type="retinoblastoma"					
	/clone_lib="Y79AAL"					
	/note="cloning vector: pME18SFL3"					
	47..1912					
	/note="unnamed protein product"					
	/codon_start=1					
	/protein_id="BAB14775.1"					
	/db_xref="GI:10436258"					
	/translation="MSGGGLFRTTAAARGLVSTANRLRLTSPVPRAFAKELF LGKIKKEVFPPEVSODELNEINQFLGPVEKFFTEEDSRKIDQEKIPDETLEKLF SLGLFQVPEEYGLGFSNTMYSLRGEIISWDGSIITVLAHQAIGLUGIILAGTPE QAKYLPKLSGEHIAAFLCTEPASGDAASIRSRATLSEDKHYILNGSKVITNGG LANITVFATKEVSDGSKVDKTAFTVERDQGVNKGKPEDKLIGKSTCEVHFE NTKIPVENLGEVDGFKVANNILNSRFSMGVSVAGLLKRLIEATYARKOFNE RLSERGLQEFALMAQRAYVHESMYTLTGMDDQPPDCIEAAMKYFSSAAWQ CVSEALQILGGIGTRDYPRIILRDLTRILFIFENELRMIALTGLQAHGRILIT RIHELQAKVSTVMDYVGRRLRSLGRVTLGLTGNHGVHPHSLADSANKPEENTYCF GRVETLLLRFGKTIIMEQLVKRVLINILYGMTAVLSRSRSIRIGLRNHDHVEL LANITFCVEAYLQNLFSLSOLDKYAPENLDEIQIKVSQILEKRAYICAHPLDRYC"					
BASE COUNT	597 a	606 c	672 g	524 t		
ORIGIN						
Query Match	99.8%; Score 1859.8; DB 9; Length 2399;					
Best Local Similarity	99.9%; Pred. No. 0;					
Matches 1861; Conservative	0; Mismatches 2; Indels 0; Gaps 0;					
QY	1	ATGAGCGGCTCGGGCTCTTCTCGGACACAGCGTGGGCTCGGCTCGGCTCGGCGGGTCTG	60			
DB	47	ATGAGCGGCTCGGGCTCTTCTCGGACACAGCGTGGGCTCGGCTCGGCTCGGCGGGTCTG	106			
QY	61	GTGCTCTCTACCGCGGAACCGCGGCTACTCGGCACACAGCGCGCTGTACGAGCTTTCGCC	120			

Db	1187	ATGGTGAAGGCTTTCAGCTCCGAGCGCGCTGAGTGTGTGAGTGAGCGCGCTGCAGATC	1246
QY	1201	CTCGGGGGCTTGGGCTACACAGGAGGACATCCGTACGAGCGATACCTCGTGCACACCGC	1260
Db	1247	CTCGGGGGCTTGGGCTACACAGGAGGACATCCGTACGAGCGATATGCGTGACACCGC	1306
QY	1261	ATCCCTCTCATCTTCGAGGGAACCAATGAGATTCCTCCGATGTACATCGCCCTGACGGT	1320
Db	1307	ATCCCTCTCATCTTCGAGGGAACCAATGAGATTCCTCCGATGTACATCGCCCTGACGGT	1366
QY	1321	CTGCAGATGCGCGCGGACATCTCTGACTACCGAGATCCATGAGCTTAACAGGCCAAAGTG	1380
Db	1367	CTGCAGATGCGCGCGGACATCTCTGACTACCGAGATCCATGAGCTTAACAGGCCAAAGTG	1426
QY	1381	AGCACAGTATGATACCTGCTTCGCGGAGGCTTCGGGACTCCCTGGCGGCACTGTGGAC	1440
Db	1427	AGCACAGTATGATACCTGCTTCGCGGAGGCTTCGGGACTCCCTGGCGGCACTGTGGAC	1486
QY	1441	CTGGGCTGACAGGCAACCATGAGTTGTGCACCCAGCTTTGCGGACAGTGCACAAAG	1500
Db	1487	CTGGGCTGACAGGCAACCATGAGTTGTGCACCCAGCTTTGCGGACAGTGCACAAAG	1546
QY	1501	TTTGAGGAGAACCTACTGCTTCGCGGAGGCTTCGGGACTCCCTGGCGGCACTGTGGAC	1560
Db	1547	TTTGAGGAGAACCTACTGCTTCGCGGAGGCTTCGGGACTCCCTGGCGGCACTGTGGAC	1606
QY	1561	AAGACCATCATGAGGAGCAGCTGCTTCTGGGCAACACTTCTCGTGGAAGCTTACCTGAGATCTC	1620
Db	1607	AAGACCATCATGAGGAGCAGCTGCTTCTGGGCAACACTTCTCGTGGAAGCTTACCTGAGATCTC	1666
QY	1621	TATGGCATGACGGCGCTGCTTCGCGGAGGCTTCGCGGACTCCCTGGCGGCACTGTGGAC	1680
Db	1667	TATGGCATGACGGCGCTGCTTCGCGGAGGCTTCGCGGACTCCCTGGCGGCACTGTGGAC	1726
QY	1681	CACGACACGAGGCTTCTTCTGGGCAACACTTCTCGTGGAAGCTTACCTGAGATCTC	1740
Db	1727	CACGACACGAGGCTTCTTCTGGGCAACACTTCTCGTGGAAGCTTACCTGAGATCTC	1786
QY	1741	TTCAGGCTCTCTCAGCTGGACAAAGTATGCTCCAGAAACCTAGATGAGCAGATTAGAA	1800
Db	1787	TTCAGGCTCTCTCAGCTGGACAAAGTATGCTCCAGAAACCTAGATGAGCAGATTAGAA	1846
QY	1801	GTGTCACAGACATCTTCGAGAGGAGGCTTATCTGTGCCACCTCTGGACAGGACA	1860
Db	1847	GTGTCACAGACATCTTCGAGAGGAGGCTTATCTGTGCCACCTCTGGACAGGACA	1906
QY	1861	TGC 1863	
Db	1907	TGC 1909	
RESULT 6			
BC007970			
LOCUS			
DEFINITION	BC007970	2494 bp mRNA linear PRI 20-MAY-2003	
ACCESSION			
VERSION	BC007970.1	GI:14044101	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalski, O., Small, D.E., Schneringer, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 2494)

Strausberg, R.

Direct Submission

Submitted (11-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@nih.gov

Akter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghghi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL plate: 20 Row: p Column: 20

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

source

1. .2494

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14452 IMAGE:4304209"

/tissue_type="uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_46"

/lab_host="DH10B-R"

/notes="Vector: pOTB7"

1. .2494

/gene="VLCAD"

/note="synonyms: NPD002, MGC14452, ACAD-9"

/db_xref="LocusID:28976"

89. .1954

/codon_start=1

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

/protein_id="AA047970.1"

/db_xref="GI:14044102"

/db_xref="LocusID:28976"

/translation="MSGCGLFRLTTAAACRACGLVSTANRLRLTSPVPAFAKELF LGIKKKEVFPFVPEVQDELNEIQFLGPEKFFTEVDSRKLQDQKIPDELEKLG SLGLFGLQVPEYGLGFSNTMYSLRGEIISMDGSIIVTAAHQALGLKGLIAGTEE OKAKYLPKLAGSHIAAFCLETPASGDAASIRSRATLSDEKKHYTLNGSKVITNGG LANIFVFAKTEVVDGSGVKDKITAFIVDERDFGGVTKGPKGLGRGNTCEVHFE NTKIPVENTILGEVDGFKVAMNINLSGRFSMGVSVAGLLKRLIEMTAEXACTRQKFNK

BASE COUNT 643 a 621 c 692 g 538 t

[illegible]

RESULT 7
AF327351
LOCUS
DEFINITION


```
ACCESSION AF327351
VERSION AF327351.1 GI:18028282
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2411)
AUTHORS Zhang, J., Zhang, W., Zou, D., Chen, G., Wan, T., Zhang, M. and Cao, X.
TITLE Cloning and functional characterization of ACAD-9, a novel member
of human acyl-CoA dehydrogenase family
JOURNAL Biochem. Biophys. Res. Commun. 297 (4), 1033-1042 (2002)
MEDLINE 2246103
PUBMED 12359260
REFERENCE 2 (bases 1 to 2411)
AUTHORS Zhang, W., Zhang, J., Li, N., Wan, T., Chen, T., Zhang, M. and Cao, X.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2000) Department of Immunology, Second Military
Medical University & Shanghai Brilliance Biotechnology Institute,
800 Xiangyin Rd., Shanghai 200433, P.R. China
LOCATION/Qualifiers
FEATURES
source
1..2411
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
33..1898
/codon_start=1
/product="very-long-chain acyl-CoA dehydrogenase VLCAD"
/protein_id="AAL56011.1"
/db_xref="GI:18028283"
/translation="MSGCGLFRTTAAARACRGLVYVSTANRLRLTSPVRAFAKELF
LGKIKKEVFPPEPVSQDELNEINQFLGVERKFTFEVDSRKIDQEKTPDETLEKLL
SLGLFQLOPEVGGIGFSNTWYSLRGLIISMDGSIITVLAHQAHLGLKILLAGTEE
QKAKYLPKLSAHEHIAFCITSPAGSDAASIRSRATLSDDKXHYLNGSKVWITNGG
LANIFVFAKTEVVDSDGSVKITAFIYVERDEGVVNGKPKDKLGRSNTCEVHFE
NTKIPENILGEVDFKVMNINLSRFSMGDSVYAGLKLRLIEMTAFAEYACRKFQFNK
RISEFLIOEKFAKMAOKAYVMESMTYLTAGMDQDPFPCDSEAAWVAFVSERAWQ
CVSEVLILGGIGYTDYPIRILDRITLLIPESNEILRMVYIALTQLOHAGRILTT
RIHELQAKVSTVMDTVGRRLRSLGRTVDLGTGNHGVHFSLADSNKFFEITYCF
GRVETFLLRFGKTIEMEOVLKRVANILINLYGMTAVLSRASRSIRIGLRNHDHEVL
LANFVCEAYLQNLFSLSQLDKYAPENLDEQIKKVSQILEKRAYICAHPLDRTC"
BASE COUNT 612 a 604 c 669 g 526 t
ORIGIN
Query Match 99.7%; Score 1858.2; DB 9; Length 2411;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1860; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ATGAGCGGCTGCGGGCTCTTCCTGCGCACACGCGTGGCGCTGCTGCGCTGCGGGGTCTG 60
Db 1 ATGAGCGGCTGCGGGCTCTTCCTGCGCACACGCGTGGCGCTGCTGCGCTGCGGGGTCTG 92
Qy 61 GTGCTCTACCGGGAACCGCGGCTACTCGGCACACGCGCGCTGTACGAGCTTTGCGC 120
Db 61 GTGCTCTACCGGGAACCGCGGCTACTCGGCACACGCGCGCTGTACGAGCTTTGCGC 152
Qy 121 AAAGAGCTTTTCTTAGCAAAAATCAGAAGAAAGAGTTTCCCATTTCCAGAGTTAGC 180
Db 121 AAAGAGCTTTTCTTAGCAAAAATCAGAAGAAAGAGTTTCCCATTTCCAGAGTTAGC 212
Qy 181 CAAGATGACTTAATGAATCAATCACTTCTTGGACCCCGTGGAAAAATTTCTCACTGAA 240
Db 213 CAAGATGACTTAATGAATCAATCACTTCTTGGACCCCGTGGAAAAATTTCTCACTGAA 272
Qy 241 GAGGTGGAATCCCGAAAAATGACAGAGGAAAAATCCAGATGAATGAAATTTGGAGAAA 300
Db 273 GAGGTGGAATCCCGAAAAATGACAGAGGAAAAATCCAGATGAATGAAATTTGGAGAAA 332
Qy 301 TTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCCAAGAGATATGCTGGCTGGGCTC 360
Db 333 TTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCCAAGAGATATGCTGGGCTGGGCTC 392
Qy 361 TCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCATCCTGTG 420
393 TCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCATCCTGTG 452
421 ACCCTGCGAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTTGGCTGGCCTAGGAG 480
453 ACCCTGCGAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTTGGCTGGCCTAGGAG 512
481 CAGAAAGCAAAATACTTTGCCCTAAACTGGCGTCCGGGAGACACATTGAGCCCTTCTGCTC 540
513 CAGAAAGCAAAATACTTTGCCCTAAACTGGCGTCCGGGAGACACATTGAGCCCTTCTGCTC 572
541 ACGGAGCCAGCCAGTGGGAGCGATGCGAGCCTCAATCCGGAGCAGAGCCACACTAAGTGAA 600
573 ACGGAGCCAGCCAGTGGGAGCGATGCGAGCCTCAATCCGGAGCAGAGCCACACTAAGTGAA 632
601 GACAAAGAGCACTACATCTCTCAATGGCTCCAAAGCTCTGGATTACTTAATGGAGGACTGGCC 660
633 GACAAAGAGCACTACATCTCTCAATGGCTCCAAAGCTCTGGATTACTTAATGGAGGACTGGCC 692
661 AATATTTTACTGTGTTTGGCAAGACATGAGTCGTTGATTCTGATGGATCAGTGAAGAC 720
693 AATATTTTACTGTGTTTGGCAAGACATGAGTCGTTGATTCTGATGGATCAGTGAAGAC 752
721 AAAATCACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTTAATGGGAAACCCGAA 780
753 AAAATCACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTTAATGGGAAACCCGAA 812
781 GATAAATTAGGCATTTCGGGGCTCCAAACACTTTGTGAAGTCCATTTTGAAGAACACCAAGATA 840
813 GATAAATTAGGCATTTCGGGGCTCCAAACACTTTGTGAAGTCCATTTTGAAGAACACCAAGATA 872
841 CCGTGTGAAAACATCTTTGGAGAGTCGAGATGGTTTAAAGTGGCCATGAACATCCTC 900
873 CCGTGTGAAAACATCTTTGGAGAGTCGAGATGGTTTAAAGTGGCCATGAACATCCTC 932
901 AACAGCGGCGGCTTCAGCATGGCGACGCTGCTGGCTGGCTGCTCAAGAGATTGATTGAA 960
933 AACAGCGGCGGCTTCAGCATGGCGACGCTGCTGGCTGGCTGCTCAAGAGATTGATTGAA 992
961 ATGACTGCTGAGTAGCCCTGCACAGGAACAGTTTAAACAGAGCTCAGTCAATTTGGA 1020
993 ATGACTGCTGAGTAGCCCTGCACAGGAACAGTTTAAACAGAGCTCAGTCAATTTGGA 1052
1021 TTGATTTCAGGAGAAATTTTGCATGTGCTCAGAGGCTTACGTCTATGAGAGTATGACC 1080
1053 TTGATTTCAGGAGAAATTTTGCATGTGCTCAGAGGCTTACGTCTATGAGAGTATGACC 1112
1081 TACCTCACAGCAGGAGTGTGGACCAACTGGCTTTCCGACTGCTCCATCAGGACGCC 1140
1113 TACCTCACAGCAGGAGTGTGGACCAACTGGCTTTCCGACTGCTCCATCAGGACGCC 1172
1141 ATGGTCAAGGTGTTACGCTCCGAGGCGCTTGGCAGTGTGAGTGAGGCTGACAGATC 1200
1173 ATGGTCAAGGTGTTACGCTCCGAGGCGCTTGGCAGTGTGAGTGAGGCTGACAGATC 1232
1201 CTGCGGGGCTTGGGCTTACACAGGGGACTATCCGTACGAGCGCATCTGGTGGACACCCGC 1260
1233 CTGCGGGGCTTGGGCTTACACAGGGGACTATCCGTACGAGCGCATCTGGTGGACACCCGC 1292
1261 ATCCTCCTCATCTTCAGGGAACCAATGAGATTTCCGGATGTACATCCCTTGACGGGT 1320
1293 ATCCTCCTCATCTTCAGGGAACCAATGAGATTTCCGGATGTACATCCCTTGACGGGT 1352
1321 CTCAGCATGCCGCGCATCTCTGACTTACAGGATCCATGAGCTTAAACAGGCAAAAGTG 1380
1353 CTCAGCATGCCGCGCATCTCTGACTTACAGGATCCATGAGCTTAAACAGGCAAAAGTG 1412
1381 AGCACAGTTCATGATACCGTTGGCGGAGGCTTCGGAGCTCCCTGGGCCCCAAGTGGAC 1440
1413 AGCACAGTTCATGATACCGTTGGCGGAGGCTTCGGAGCTTCGGGCGGAACTGGGAC 1472
1441 CTGGGCTGACAGGCAACCATGAGTTGTGACCCAGTCTTGGGACAGTGGCAACAAG 1500
```

Db 1473 CTGGGGTGACAGGCAACACCTGGAGTTGTGCACCCAGTCTTGGGACAGTGCCACAAG 1532
 Qy 1501 TTTGAGGAGAACACTACTGCTTCGGCCGGACCGTGGAGACACTGCTGCTCCGCTTTGGC 1560
 Db 1533 TTTGAGGAGAACACTACTGCTTCGGCCGGACCGTGGAGACACTGCTGCTCCGCTTTGGC 1592
 Qy 1561 AAGACCATCATGAGGAGCAGCTGGTACTGAAGCGGTGGCCAAACATCTCATCAACCTG 1620
 Db 1593 AAGACCATCATGAGGAGCAGCTGGTACTGAAGCGGTGGCCAAACATCTCATCAACCTG 1652
 Qy 1621 TATGGCATGACGGCGTCTGCTCGCGGCCAGCGCTCCATCGGCATTTGGGCTCCGCAAC 1680
 Db 1653 TATGGCATGACGGCGTCTGCTCGCGGCCAGCGCTCCATCGGCATTTGGGCTCCGCAAC 1712
 Qy 1681 CAGCACACAGAGTCTCTTGGCCCAACACCTTCTGCGTGGAACTTACTTGCAGATCTC 1740
 Db 1713 CAGCACACAGAGTCTCTTGGCCCAACACCTTCTGCGTGGAACTTACTTGCAGATCTC 1772
 Qy 1741 TTCAGCCTCTCAGCTGGACAAGTATGTCTCCAGAAAACCTAGATGAGCAGATTAAGAAA 1800
 Db 1773 TTCAGCCTCTCAGCTGGACAAGTATGTCTCCAGAAAACCTAGATGAGCAGATTAAGAAA 1832
 Qy 1801 GTGTCCAGCAGATCTCTTGAGAAGCGAGCTATATCTGTGCCACCTCTGGACAGGACA 1860
 Db 1833 GTGTCCAGCAGATCTCTTGAGAAGCGAGCTATATCTGTGCCACCTCTGGACAGGACA 1892
 Qy 1861 TGC 1863
 Db 1893 TGC 1895

RESULT 8

BD157796 2400 bp DNA linear PAT 17-JAN-2003
 LOCUS Primer for synthesizing full-length cDNA and use thereof.
 DEFINITION BD157796
 ACCESSION BD157796
 VERSION BD157796.1 GI:27863554
 KEYWORDS JP 2002191363-A/12639

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2400)
 Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
 Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 12639 09-JUL-2002;
 HELIX RESEARCH INSTITUTE

COMMENT

OS Homo sapiens (human)
 PN JP 2002191363-A/12639
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
 PI SAITO,
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KETICHI NAGAI, TETSUJI OTSUKI
 PC
 C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
 10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key

FEATURES

FT CDS Location/Qualifiers

1..2400

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT 570 a 629 c 685 g 516 t

ORIGIN

Query Match

Best Local Similarity 89.7%; Score 1671.4; DB 6; Length 2400;

Pred. No. 0;

Matches 1807; Conservative 0; Mismatches 56; Indels 82; Gaps 2;
 Qy 1 ATGAGCGGTGGGGGCTCTTCTTGCACACACGGCTCGGCTCGTCCCTCCGGGCTCTG 60
 Db 47 ATGAGCGGTGGGGGCTCTTCTTGCACACACGGCTCGGCTCGTCCCTCCGGGCTCTG 106
 Qy 61 GTGGTCTCTACCGCAACCGGGGCTTACTGGCCACAGCCCGCTGTAGAGCTTTTCGCC 120
 Db 107 GTGGTCTCTACCGCAACCGGGGCTTACTGGCCACAGCCCGCTGTAGAGCTTTTCGCC 166
 Qy 121 AAGAGCTTTTCTTAGGCAAAATCAAGAAGAAA----- 153
 Db 167 AAGAGCTTTTCTTAGGCAAAATCAAGAAGAAA----- 226
 Qy 154 -----GAAGTTTTCCATTTTCCAG 172
 Db 227 TCTGGTCCCGCTTTTTCACCCCTCAGCTGCAAGACTGGTGTGAACCTTTGTGAGATTCCCC 286
 Qy 173 AAGTTAGCCCAAGATGA-----ACTTAATGAATCAATCAGTTCTTGGGACC 218
 Db 287 AACCTGCCAGAGAGATACACCTGGGCGGAGGGGTGTTAAACCTCCGGATTTCTGAGT 346
 Qy 219 CBTGAAAATTTCTTCACTGAAGAGTGGACTCCGAAAAATTTGACAGGAAGGAAAAAT 278
 Db 347 TCCAGAAAACCTTCCAGAGAAAGGTGGACTCCCGAAAAATTTGACAGGAAGGAAAAAT 406
 Qy 279 CCCAGATGAACCTTTGGAGAAATTTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCCA 338
 Db 407 CCCAGATGAACCTTTGGAGAAATTTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCCA 466
 Qy 339 AGAATATGGTGGCTTCTTCCAAACACCATGTACTCAAGACTTAGGGAGATCATCAG 398
 Db 467 AGAATATGGTGGCTTCTTCCAAACACCATGTACTCAAGACTTAGGGAGATCATCAG 526
 Qy 399 CATGATGGTCCATCACTGTGACCTGGCAGCCAGCAGGCTATTGGCTCAAGGGAT 458
 Db 527 CATGATGGTCCATCACTGTGACCTGGCAGCCAGCAGGCTATTGGCTCAAGGGAT 586
 Qy 459 CATCTTGGCTGGCACTGAGGAGCAAGCAAAATCTTGCTTAACCTGGGCTCCGGGGA 518
 Db 587 CATCTTGGCTGGCACTGAGGAGCAAGCAAAATCTTGCTTAACCTGGGCTCCGGGGA 646
 Qy 519 GCACATTCAGCCCTTCTGCTTCCAGGAGCCAGCCAGTGGGAGCGATGAGGCTCAATCCG 578
 Db 647 GCACATTCAGCCCTTCTGCTTCCAGGAGCCAGCCAGTGGGAGCGATGAGGCTCAATCCG 706
 Qy 579 GAGCAGCCACACTAAGTGAAGCAAGCAAGCACTACATCTCAATGGCTCCAGGCTG 638
 Db 707 GAGCAGCCACACTAAGTGAAGCAAGCAAGCACTACATCTCAATGGCTCCAGGCTG 766
 Qy 639 GATTACTAATGGAGCTGGCCAAATATTTTACTGTGTGTTTCAAGACTGAGGCTCGTTGA 698
 Db 767 GATTACTAATGGAGCTGGCCAAATATTTTACTGTGTGTTTCAAGACTGAGGCTCGTTGA 826
 Qy 699 TTTGATGGATCAGTGAAGCAAAATCACAGCATTCATAGTAGAAAAGACTTTGGTGG 758
 Db 827 TTTGATGGATCAGTGAAGCAAAATCACAGCATTCATAGTAGAAAAGACTTTGGTGG 886
 Qy 759 AGTCACATAATGGGAAACCCGGAAGATAAATTAAGGCATTTCCGGGCTCCAAACACTTTG 818
 Db 887 AGTCACATAATGGGAAACCCGGAAGATAAATTAAGGCATTTCCGGGCTCCAAACACTTTG 946
 Qy 819 CCATTTTGAAGAACCAAGATACCTGTGGAAGAACATCTTTGGAGAGCTCGGAGATGGGTT 878
 Db 947 CCATTTTGAAGAACCAAGATACCTGTGGAAGAACATCTTTGGAGAGCTCGGAGATGGGTT 1006
 Qy 879 TAAGGTGCCATGAACATCTTCAACAGCGCGGTTTCCAGCATGGGCAAGCTTGGCTGG 938
 Db 1007 TAAGGTGCCATGAACATCTTCAACAGCGCGGTTTCCAGCATGGGCAAGCTTGGCTGG 1066
 Qy 939 GCTCCTCAAGAGATTGATTGAATGACTGCTGAGTAGCCCTGCACAGGAAGAACTTTAA 998
 Db 1067 GCTCCTCAAGAGATTGATTGAATGACTGCTGAGTAGCCCTGCACAGGAAGAACTTTAA 1126

```
QY 999 CAAGAGGCTCAGTGAATTTGGATTGATTCAGAGAAATTTGACATGATGGCTCAGAAGGC 1058
Db 1127 CAAGAGGCTCAGTGAATTTGGATTGATTCAGAGAAATTTGACATGATGGCTCAGAAGGC 1186
QY 1059 TTACGTCATGGAGAGTATGACCTACCTCAGCAGGAGTATGCTGGACCAACCTTGGCTTTC 1118
Db 1187 TTACGTCATGGAGAGTATGACCTACCTCAGCAGGAGTATGCTGGACCAACCTTGGCTTTC 1246
QY 1119 CGACTCTCATCGAGGAGCCGANTGTTGAAGTGTTCAGCTCCGAGGCCGCTGGCAGTG 1178
Db 1247 CGACTCTCATCGAGGAGCCGANTGTTGAAGTGTTCAGCTCCGAGGCCGCTGGCAGTG 1306
QY 1179 TGTGAGTGAAGGCTCAGATCTCGGGGCTTGGCTACACAAAGGACTATCCGTACGA 1238
Db 1307 TGTGAGTGAAGGCTCAGATCTCGGGGCTTGGCTACACAAAGGACTATCCGTACGA 1366
QY 1239 GCGCATACGCTGACACCCGATCCTCCTCATCTTCGAGGGAACCAATGAGATTCGCG 1298
Db 1367 GCGCATACGCTGACACCCGATCCTCCTCATCTTCGAGGGAACCAATGAGATTCGCG 1426
QY 1299 GATGTACATCGCCCTGACGGTCTGCAGATCGCGCGCATCCTCACTACAGGATCCA 1358
Db 1427 GATGTACATCGCCCTGACGGTCTGCAGATCGCGCGCATCCTCACTACAGGATCCA 1486
QY 1359 TGAGCTTAACAGAGCCAAAGTAGACACAGTATGATACCGTTGGCCGAGGCTTCGGGA 1418
Db 1487 TGAGCTTAACAGAGCCAAAGTAGACACAGTATGATACCGTTGGCCGAGGCTTCGGGA 1546
QY 1419 CTCCTGGGCGCAACTGTGACCTGGGCTGACAGCAACCATGAGTGTGCAACCCAG 1478
Db 1547 CTCCTGGGCGCAACTGTGACCTGGGCTGACAGCAACCATGAGTGTGCAACCCAG 1606
QY 1479 TCTTGGGACAGTCCCAAGTTTGAGGAGAACCTACTGCTTCGGCGGACCGTGA 1538
Db 1607 TCTTGGGACAGTCCCAAGTTTGAGGAGAACCTACTGCTTCGGCGGACCGTGA 1666
QY 1539 GACACTGCTCCTCGCTTTGGCAAGACCATCATGAGGAGGACGCTGGTACTGAAGCGGT 1598
Db 1667 GACACTGCTCCTCGCTTTGGCAAGACCATCATGAGGAGGACGCTGGTACTGAAGCGGT 1726
QY 1599 GGCACACATCCTCATCAACTGTATGCGATGACGGCGTCTCTCGGGGCCAGCGCTC 1658
Db 1727 GGCACACATCCTCATCAACTGTATGCGATGACGGCGTCTCTCGGGGCCAGCGCTC 1786
QY 1659 CATCCGATTTGGGCTCGGCAACACGACGAGGTTCTCTTGGCCAAACCTTCTGGGT 1718
Db 1787 CATCCGATTTGGGCTCGGCAACACGACGAGGTTCTCTTGGCCAAACCTTCTGGGT 1846
QY 1719 GGAAGCTTACTGAGAAATCTTTCAGCTCTCTCAGCTGGAAGTATGCTCCAGAAA 1778
Db 1847 GGAAGCTTACTGAGAAATCTTTCAGCTCTCTCAGCTGGAAGTATGCTCCAGAAA 1906
QY 1779 CTTAGATGAGCAGATTGAAGAAAGTGTCCAGCAGATCCTTGAAGAGGAGCCTATATCTG 1838
Db 1907 CTTAGATGAGCAGATTGAAGAAAGTGTCCAGCAGATCCTTGAAGAGGAGCCTATATCTG 1966
QY 1839 TGCCCCACCTCTGTGACAGGACATGC 1863
Db 1967 TGCCCCACCTCTGTGACAGGACATGC 1991
```

```
RESULT 9
AK022568
LOCUS Homo sapiens cDNA FLJ12506 fis, clone NT2RM2001700, weakly similar
DEFINITION to ACYL-COA DEHYDROGENASE, VERY-LONG-CHAIN SPECIFIC (EC 1.3.99.-).
ACCESSION AK022568
VERSION AK022568.1 GI:10434032
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE
AUTHORSTITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES
Source

CDS

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1
Isoqai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
Nishikawa, T., Nagai, K., Sugano, S., Aotsuka, S., Yoshikawa, Y.,
Matsumura, H., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,
Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y. and Sasaki, N.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2400)
Isoqai, T. and Otsuki, T.
Direct Submission
Submitted (23-AUG-2000) Takao Isoqai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- and 3'-end one pass sequencing and clone selection;
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
Location/Qualifiers
1..2400
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RM2001700"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RM2"
/note="cloning vector: pME18SFL3-mRNA from uninduced NT2
neuronal precursor cells."
498..1994
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14104.1"
/db_xref="GI:10434032"
/translation="MYSRLGEIISMDGSTVTTLAAHOAIGLKGILLAGTEOKAKYLP
KLAGSHIAAFLTEPASGSDASTRSATLSLEDKHYILNGSKVMTTNGLIATFTV
FAKTEVVDSDGSKDKITAFIVERDFGVTNGKPKDKLGIRNSNCEVFNKIPVE
NLTGEVDFGSKVAMNINSGFRSMGVSAGLRLLEMTAEIACIRKQNKRLSEFGL
IOEKFALMAQKAYVMESMTYLAGMDQDFPDCSEAAVMKVFSSEAAQVCSEALQ
ILGGLGYRDYPERILTRILLIPEGTNEILRMVIALTLQHGAGRLITTRITHELKQ
AKVSTVMDTVGRRLRLGTLGTLGNHGVHPSLADSAKFENFTYCFGRVETL
LIRFCKTLMEEOLVLKRVANILNLVMTAVLSRASKSTRIGLRNHDHVLANTFCV
EAYLQNLFSLSQLDKYAPENLDEQIKKVSQQLLEKRAYICAHPLDRTC"

BASE COUNT 570 a 629 c 685 g 516 t
ORIGINQuery Match 89.7%; Score 1671.4; DB 9; Length 2400;
Best Local Similarity 92.9%; Pred. No. 0;
Matches 1807; Conservative 0; Mismatches 56; Indels 82; Gaps 2;

```
QY 1 ATGAGCGGCTGCGGGCTCTTCTCGCGCACCCACCGGCTCGGCTCGCTCGCTCGCGGGCTCG 60
Db 47 ATGAGCGGCTGCGGGCTCTTCTCGCGCACCCACCGGCTCGGCTCGCTCGCTCGCGGGCTCG 106
QY 61 GTGGTCTCTACCGCGAAGCGGCGGTCTACTGCGCACCGCGCGCTGTACGAGCTTTCGCC 120
Db 107 GTGGTCTCTACCGCGAAGCGGCGGTCTACTGCGCACCGCGCGCTGTACGAGCTTTCGCC 166
QY 121 AAAGAGCTTTTCTTAGGCAAAATCAAGAAGAAA----- 153
Db 167 AAAGAGCTTTTCTTAGGCAAAATCAAGAAGAAA----- 226
QY 154 ----- 172
Db 227 TCTGGCTCCGCTTTTACCCCTCAGCTGCAAGACTGGTGTGAATTTGTGAGATTCGCC 286
QY 173 AAGTTAGCCCAAGATGA-----ACTTAATGAATCAATCAGTCTTGGGACC 218
Db 287 AAACCTGCCAGAGAGATACACCCTGGCGCCGAGCGGTGTTAACACTCCCGATTCTCGT 346
```

QY	219	CGTGGAAAAATTCCTCAGTGAAGAGGTGGACTCCCGAAAAATTTGACAGGAAGGAAAT	278
Db	347	TCACAGAAACCTTTCCAGAGAAAGGTGACTCCCGAAAAATTTGACAGGAAGGAAAT	406
QY	279	CCAGATGAACATTTTGAGAGAAATGAAGAGCCTTAGGCGTTTGGGCTCCAAAGTCCCGAGA	338
Db	407	CCAGATGAACATTTTGAGAGAAATGAAGAGCCTTAGGCGTTTGGGCTCCAAAGTCCCGAGA	466
QY	339	AGAATATGGTGGCTGGGCTTCCTCAACACCATGTACTCAAGACTAGGGAGATCATCAG	398
Db	467	AGAATATGGTGGCTGGGCTTCCTCAACACCATGTACTCAAGACTAGGGAGATCATCAG	526
QY	399	CATGATGGGTCCATCACTGTGACCCCTGGGAGCCGACAGGCTATTGGCTCAAGGGGAT	458
Db	527	CATGATGGGTCCATCACTGTGACCCCTGGGAGCCGACAGGCTATTGGCTCAAGGGGAT	586
QY	459	CATCTTGGCTGGCACTGAGAGCAGAGAAAGCAAAATACCTTAACCTAGGCGTCCGGGGA	518
Db	587	CATCTTGGCTGGCACTGAGAGCAGAGAAAGCAAAATACCTTAACCTAGGCGTCCGGGGA	646
QY	519	GCACATTTGCAGCCTTTCCTCAGGAGCAGCCAGTGGGAGCGATGAGCCTCAATCCG	578
Db	647	GCACATTTGCAGCCTTTCCTCAGGAGCAGCCAGTGGGAGCGATGAGCCTCAATCCG	706
QY	579	GAGCAGAGCACACTAAGTGAAGACAGAGCACTACATCCCTCAATGGCTCCAAAGTCTG	638
Db	707	GAGCAGAGCACACTAAGTGAAGACAGAGCACTACATCCCTCAATGGCTCCAAAGTCTG	766
QY	639	GATTACTAATGGAGGACTGGCCAAATATTTTACTGTGTTTGCAGAGACTGAGTCTGTTGA	698
Db	767	GATTACTAATGGAGGACTGGCCAAATATTTTACTGTGTTTGCAGAGACTGAGTCTGTTGA	826
QY	599	TTCTGATGATCAGTGAAGACAAAAATCAGCATTTCATAGTAGAAAGAGACTTTGGTGG	758
Db	827	TTCTGATGATCAGTGAAGACAAAAATCAGCATTTCATAGTAGAAAGAGACTTTGGTGG	886
QY	759	AGTCACCTAATGGGAACCCAGAGATAAATAGGCAATTCGGGCTCCAAACACTTTGTGAAGT	818
Db	887	AGTCACCTAATGGGAACCCAGAGATAAATAGGCAATTCGGGCTCCAAACACTTTGTGAAGT	946
QY	819	CCATTTTGAACACCAAGATACCTGTGGAACCAATCTTGGAGAGGTCCGAGATGGGTT	878
Db	947	CCATTTTGAACACCAAGATACCTGTGGAACCAATCTTGGAGAGGTCCGAGATGGGTT	1006
QY	879	TAAGGTGGCATGAACATCCTCAACAGCGCGGTTTCAGCTAGGCGAGCTCGTGGCTGG	938
Db	1007	TAAGGTGGCATGAACATCCTCAACAGCGCGGTTTCAGCTAGGCGAGCTCGTGGCTGG	1066
QY	939	GCTGCTCAAGAGATTGATTGAATGACTGCTGAGTACCCCTGCAGAGAAACAGTTTAA	998
Db	1067	GCTGCTCAAGAGATTGATTGAATGACTGCTGAGTACCCCTGCAGAGAAACAGTTTAA	1126
QY	999	CAAGAGGCTCAGTGAATTTGGATTGATTTCAGAGAAATTTGCACTGATGCTCAGAGGC	1058
Db	1127	CAAGAGGCTCAGTGAATTTGGATTGATTTCAGAGAAATTTGCACTGATGCTCAGAGGC	1186
QY	1059	TTACGTCATGGAGATGATGACTACCTCAGCAGAGGATGCTGGACCAACCTGGCTTTC	1118
Db	1187	TTACGTCATGGAGATGATGACTACCTCAGCAGAGGATGCTGGACCAACCTGGCTTTC	1246
QY	1119	CGACTGCTCATTCAGGAGCAGCATGGTGAAGGTGTTTCAGCTCCGAGGCCCTCGCAGTG	1178
Db	1247	CGACTGCTCATTCAGGAGCAGCATGGTGAAGGTGTTTCAGCTCCGAGGCCCTCGCAGTG	1306
QY	1179	TGTGAGTGAAGGCTGCAGATCCTCGGGGCTTGGGCTACACAGGAGACTATCCGTACGA	1238
Db	1307	TGTGAGTGAAGGCTGCAGATCCTCGGGGCTTGGGCTACACAGGAGACTATCCGTACGA	1366
QY	1239	GGGCATCTGGTGACACCCCGATCTCTCCTCATCTTCGAGGGGAACCAATGAGATTCCTCG	1298
Db	1367	GGGCATCTGGTGACACCCCGATCTCTCCTCATCTTCGAGGGGAACCAATGAGATTCCTCG	1426
QY	1299	GATGTACATGCGCTGACGGGTCTGCGATGCTGGCGGCGCATCTCTGACTACAGAGATCA	1358

JOURNAL MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	Db	74	CGGCTGGCGGCGAGCTTCCCGGTGCTGAGGGAATTCACGGCGAGCGCGCGCGCTGC	133
REFERENCE	2 (bases 1 to 2129)	Qy	92	GCACCGCCCGCTGTAGAGCTTTCGCAAGAGCTTTCTCTAGGCAAAATCAAGAAGA	151
AUTHORS	Strausberg, R.	Db	134	ACACCGAGCTCGAGAGCTGCTGCTGCGCAAGAGAGCTTCTCTGGGCAACATCAAGCAGA	193
JOURNAL	Submitted (26-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	Qy	152	AAGAAGTTTCCCATTTCCAGAAAGTTAGCCAAAGATGAACCTTAATGAATCAATCAGTTCT	211
REMARK COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgabps@mail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: ang@bcm.tmc.edu Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.	Db	194	AAGGAGTTTCCCATTTCCAGAGGTGAGTGAACCTTAGTGAATCAATCAGTTCTG	253
		Qy	212	TGGACCGCTGGAAAAATTTCTTCACTGAAGAGTGGAGTCCCGAAAAATTCACGAGGAAG	271
		Db	254	TGGACCGCTGGAAAAATTTCTTCACTGAAGAGTGGAGTCTCGAAAAATTCACGAGGAG	313
		Qy	272	GGAAAAATCCAGATGAACCTTTGGAGAAATTAAGAGAGCTAGGCGCTTTTGGGTGCAAG	331
		Db	314	GAANAATCCAGTTCACACCTTAGAGAAGTTGAAGAGCTGGGACTTTTGGCATACAGG	373
		Qy	332	TCCCAAGAATATGGTGGCGCTTCTCAACACCATCTACTCAAGACTAGGGGAGA	391
		Db	374	TCCCAAGAATATGGTGGCGCTTCTCAACACCATCTACTCGGCTCGGGGAGA	433
		Qy	392	TCATCAGCATGATGGTCCATCACTGTGACCTGGCAGCGCAGCAGCTATTGCGCTCA	451
		Db	434	TCATCAGCTGGATGCCCTCTATCAGGTGACCTGGCAGCAGCAGCTATCGGCTCA	493
		Qy	452	AGGGATCATCTTGGCTGGCAGTGGAGAGCAAGCAAAATCTTAACTGGCGGT	511
		Db	494	AGGGATTTATCTTGGTGGCAATGAGCAAGAGAGGCGCAAGTATGCGCAAACTGCTG	553
		Qy	512	CGGGGAGCACATTCAGCGCTTCTGCTCAGCGCCAGCGAGTGGGAGCGATCGAGCT	571
		Db	554	CTGGGGAACACATCGAGCGCTTCTGCTCAACAGAGCCAGCGAGTGGATGCTGCAT	613
		Qy	572	CAATCCGGAGCAGCCACACTAAGTGAAGCAAGCAAGCACTTCAATCTCAATGCTCA	631
		Db	614	CCATCCAGACAGAGCTAGCTTAACTGAAGTAAGAACTTCTCTCAATGCTCA	673
		Qy	632	AGGTCTGATTTACTATGAGGAGCTGGCAATATTTTACTGTGTTCGCAAGAGCTGAGG	691
		Db	674	AGGTTTGGATCACAATGGAGGAGTGGCCCAATATTTTACTGTGTTCGCAAAACCGAGG	733
		Qy	692	TCGTTGATTTCTGATGATCAGTGAAGCAAAATCACAGCACTTCACTAGTGAAGAGACT	751
		Db	734	TGTTGATTTCTGATGTTCTGAAACACAGCAAAATGACCGCTTCACTAGTGAAGAGACT	793
		Qy	752	TTGGTGGAGTCACTAATGGGAAACCGGAAGATAATAGGCACTTGGGCTCCACACTT	811
		Db	794	TCGGCGAATCACAATGGGAAACCGGAAGATAATAGGCACTTGGGCTCCACACT	853
		Qy	812	GTGAAGTCCATTTTGAACACACCAAGATACCTGTGAAACACATCTTGGAGAGTGGAG	871
		Db	854	GTGAGTCCATTTTGAATAATCCAGAGTGGCTGCTGGAATGCTTGGAGAGTGGAG	913
		Qy	872	ATGGGTTTAAAGTGGCCATGAACATCCTCAACAGCGCGGCTTCACTAGCTGGCAGCTG	931
		Db	914	GGGCTTTAAAGTGGCCATGAACATCCTGAACAGTGGGCAATTCAGCTGGCAGTGGT	973
		Qy	932	TGGCTGGCTGCTCAGAGATTCATTTGAATGACTCTGAGTACGCTGCACAAGGAAAC	991
		Db	974	TGGCGGGGATGCTGAAGAGAGCTGATCGAAGTACTGCTGCTGATGATGCTGTACGAGGAAAC	1033
		Qy	992	AGTTTAAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGACATGATGGCTC	1051
		Db	1034	AGTTTAAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGACATGATGGCTC	1093
		Qy	1052	AGAAGGCTTACGTCATGGAGAGTATGACCTACCTCAGCAGCGAGTGGGACCAACCTG	1111
		Db	1094	AGAAGGCTTACGTCATGGAGAGTATGACCTACCTCAGCAGCGAGTGGGACCAACCTG	1153
		Qy	1112	GCATTCGCGAGTCTCCATCGAGCAGCCATGCTGAAGGTGTTTCAAGTCCGAGCGCGCT	1171
FEATURES	Location/Qualifiers				
source	1. .2129				
	/organism="Mus musculus"				
	/mol_type="mRNA"				
	/strain="FVB/N"				
	/db_xref="taxon:10090"				
	/clone="MGC:28694 IMAGE:4241041"				
	/tissue_type="Kidney, normal. 5 month old male mouse."				
	/clone_lib="NCI_CGAP_kid4"				
	/lab_host="DH10B"				
	/note="vector: pCMV-SPORT6"				
gene	1. .2129				
	/gene="2600017P15R1k"				
	/note="synonym: 4732402K02"				
	/db_xref="LocusID:229211"				
	/db_xref="MGI:1914272"				
	31. .1908				
CDS	/codon_start=1				
	/product="RIKEN cDNA 2600017P15"				
	/protein_id="AAH33277.1"				
	/db_xref="GI:23958795"				
	/db_xref="LocusID:229211"				
	/translation="MSGCVLLSRGATAAAARASRLRFTARRRLPHTSLQSCSFA KEFLGNKQKVPPEVSHSELNPFVPLEKFTTEVDSRSDIOBKIPVDL EKLSGLGFIQVPEYGGGLNSMTWARLGEILSDASTVTAAHQALGKILV TNGQKAKYLPKLSGGHIAAFLTEPAGSDAASQTATLSEDKKYLINGSKWI TNGGLANFTVFAKTEVDSDGTDKMTAFIVERDFGGITNGKPEKIGNSNCE VFENTRPVENVLGEVGGFKVAMNINLSRFSMSAVAGMLKLIETLAEYACTRK QFNRLSEFGLIQEKFALMAQKAYVMESMAYLTSGMIDPFPDCSIEAAMVAVFSE AAMQVSEALQIOLGGGCMKDYPERMLRDARILLIPEGTNEILRLFIALTGLQHAGR ILSRILEKSGNVTVMETIGRKLRLSLGRTVDLGLTGDLGVVHPVPSLGSANKLEND VHFGRTVEILLRFGKNIVEBQILVKRVANILINLYGMTAVLSRASRSIRIGLRNHD HEVLNANMFCVEAFQNLFLSLQDKNAPENLDEQIKVSRQILEKRAYICAPLDRAS"				
BASE COUNT	554 a 495 c 586 g 494 t				
ORIGIN					
Query Match	74.1%; Score 1380.8; DB 10; Length 2129;				
Best Local Similarity	84.6%; Pred. No. 0;				
Matches 1550; Conservative	0; Mismatches 282; Indels 0; Gaps 0;				
Qy	32 CGGCTGGCGCTGCTGCGCGGTCTGCTGCTTACCGGGAACCGCGGCTACTGC				

Db 1154 GATTCCCGACTGCTCTATCGAGCGCAGCCATGCTGAAGGTGTTTAGCTCTGAGGCTGCCT 1213

QY 1172 GGCAGTGTGTGAGTGGCGCTGCAGATCCTCGGGGCTTGGCTACACAAAGGACTATC 1231

Db 1214 GGCAGTGTGTGAGGAGGCTCTGCAGATCCTTGGGGCTCGAGCTACATGAAGGACTACC 1273

QY 1232 CGTAGAGCGCATATGCTGACACCCGCGCTCTCTCATCTTCAGGGAACCAATGAGA 1291

Db 1274 CCTATGAGCGCATGCTGCTGATGCCCGCATCTCTCTATCTTTCAGGGAACCAATGAGA 1333

QY 1292 TTCTCGGATGTACATCGCCCTGACGGGTCTGCAGATCCCGCGCATCTCTGACTACCA 1351

Db 1334 TTCTTCGGTGTTCATGTGCCCTGACAGGCTCTGCAGATCTGAGCGCATCTCTGCTGCA 1393

QY 1352 GGATCCATGAGCTTAAACAGGCAAGTGCAGACAGTCAATGATACCGTTCGCCGAGGC 1411

Db 1394 GGATCAAGAGCTTAAAGTGGCAATGTGACACACAGTCAATGATGCTGCGGAAC 1453

QY 1412 TTCGGGACTCCTGGCGCAACTGTGGACCTGGGCTGACAGCAACCATGAGTGTGTC 1471

Db 1454 TTCGGGATTTCTTGGCGCAACTGTGGACCTGGGCTAACAGCGCATCTTGGTGTGTC 1513

QY 1472 ACCCGATCTTTCGGGAGTGCACAAAGTTTTCAGGAGCAACACCTACTCTCGCCGCA 1531

Db 1514 ACCCGAGCTTGGCGACAGTGCACAAAGCTTTCAGGAAATGTCATTACTTGGCCGGA 1573

QY 1532 CGGTGGAGACATGCTGCTCGCTTTGGCAAGACCATGATGAGGAGCAGCTGTGACTGA 1591

Db 1574 CTGTTGAGACCTGTTACTCCGTTTGGAAAGAACATGTAGAGGAACAGCTGGTGTAA 1633

QY 1592 AGCGGGTGGCAACATCTCATCACTGTATGGCATGACGCCCTGTGCTGCGGCGCA 1651

Db 1634 AGCGGGTGGCAACATCTCATCACTGTATGGCATGACGCCCTGTGCTGCGGAGCCA 1693

QY 1652 CGCGTCCATCGCATGCTGCTCGCTTCCCAACACGACGAGCTTCTTGGCCACACCT 1711

Db 1694 CGCGTCCATCAGATGCTGCTCGCTTCCCAACACGAGCTTCTTGGCCACACATGT 1753

QY 1712 TCTCGTGGAGCTTACTTGGCAGATCTCTTCAGCTCTCTCAGTGGACAGTATGCTC 1771

Db 1754 TCTGTTGGAGCTTATTTCCAGATCTCTTTCAGCTCTCTCAGTGGACAGAGCTC 1813

QY 1772 CAGAAACCTAGATGAGCAGATAGAAAGTGTCCACACATCTTTCAGAGCGAGCT 1831

Db 1814 CAGAAACCTGATGAGCAGATAGAAAGTGTCTGCGCAGATCTTTCAGAGCGAGCT 1873

QY 1832 ATATCTGTGCCACCTCTGAGCAGCATGC 1863

Db 1874 ATATCTGTGCCACCTCTGAGCAGCATGC 1905

RESULT 11

BC032213

LOCUS

Mus musculus RIKEN cDNA 2600017P15 gene, mRNA (cdna clone MGC:38234 IMAGE:5323723), complete cds.

ACCESSION

BC032213

VERSION

BC032213.1

KEYWORDS

MGC.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

AUTHORS

1 (bases 1 to 2148)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schreier, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Small, M.A., Schnerker, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

CONTACT: MGC help desk

Email: cgabbs@mail.nih.gov

Tissue Procurement: Jeffrey Green M.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: http://www.nisc.nih.gov/

Contact: nisc.mgc@hgrl.nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,

Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Tsurgov, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Series: IRAC Plate: 56 Row: 1 Column: 12

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis.

FEATURES

Location/Qualifiers

1..2148

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:38234 IMAGE:5323723"

/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating

ductal carcinoma. 5 month old virgin mouse."

/clone_lib="NCI_CGAP_Mam6"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1..2148

/gene="2600017P15R1K"

/note="synonym: 4732402K02"

/db_xref="LocusID:229211"

/db_xref="MGI:1914272"

51..1928

/codon_start=1

/product="RIKEN cDNA 2600017P15"

/protein_id="AAH32213.1"

/db_xref="GI:21595209"

/db_xref="LocusID:229211"

/translation="MSGCVLLSGATAAAAAASRVLRREFTARRRPLHTSLOSCSFA

KELFLGNIKQKQVFPPEVSEHSEINQFVGPLEKFFTEEDVSDRDKGKIPVDTL

EKLSLGLGFIQVPEYGGVLSNTMYARLGEITISLSDATITVTLAAHQAIGLKGITLV

GNDEQAKYLPKLSSSEHIAAFCLTEPASGSDAASIQTRATLSSEDKKYLILNGSKWVI TNGFLANITFTFAKTEVVDGSKTDKMTAFIVBRDFGCTNGKPEDKJIFLNGSNICE VHFNRTPVFNLEVGSGPKVANNILNSGFFSMGSAVAGMKLIELTAYEACTRK QNRNLSFGLIQRFAFMAQKAYVHESMAYLTSGMLDQEPFDCSIEAAMVKVFSSE AAWQCSALQILGSGGYMKDYPIYERMLRDARILIDIFEGTNEILFLFALTALQHQHAGR ILTSRIKELKSNVTVMETIGRKLRDSIGRTVDLGLDGVHPSLGDSSANKLEEN VHYFRTVETLLRFGKNIVBQLKRVANILINLYGMTAVLSRSIRIGLNHD HEVLLANMFCEAYEAFONLFSLSOLDKNAPENLDEQIKKVSQILEKRAYICAHPLDRA S*					
BASE COUNT	554 a	504 c	593 g	497 t	
ORIGIN					
Query Match	74.1%; Score 1380.8; DB 10; Length 2148;				
Best Local Similarity	84.6%; Pred. No. 0;				
Matches 1550; Conservative	0; Mismatches 282; Indels 0; Gaps 0;				
QY	32	CGGCTCGGGCTCGTGGCTGCGCGGGTCTGGTGTCTCTACCGGGAACCGCGGCTACTGCG	91		
DB	94	CGGCTCGGGCGGAGCTTCCCGGTGCTGAGGGAATTCACGGCGAGGCGCGCCGCGCTGC	153		
QY	92	GCACGACCGCGCTGTACGAGCTTTCGCCAAAGAGCTTTTCTTAGGCAAAATCAAGAGA	151		
DB	154	ACACGACGCTGCAAGAGCTGCTGCTGCCCAAGAGCTCTTCTCGGGCAACATCAAGCAGA	213		
QY	152	AAGAAGTTTTCCTATTTCCAGAAAGTTAGCCAAAGATGAATTAATGAATCAATCAGTTCT	211		
DB	214	AAGGAGTTTTCCTATTTCCAGAGGTCAGTCAACATGAATTAGTGAATCAATCAGTTCTG	273		
QY	212	TGGGACCGCTGGAAAAATTTCTCCTGAAGAGTGGACTCCGCAAAATTTGACACAGAAG	271		
DB	274	TGGGACCGCTTGGAAAAATTTCTCCTGAAGAGTGGACTCTCGAAAAATTTGACACAGAGG	333		
QY	272	GGAATAATCCAGATGAATTTTGGGAAATTTGAAGAGCTTAGGGCTTTTGGGCTGCAAG	331		
DB	334	GAAAAATCCAGTTGACACTTAGAGAGTTGAAGAGCTTGGGACTTTTGGCATACAGG	393		
QY	332	TCCGAGAAGATATGTGGCTGGGCTTCTCCAAACACCATGTACTCAAGACTAGGGGAGA	391		
DB	394	TCCGAGAAGATATGTGGCTGGGCTTCTCCAAACACCATGTATGCTCGGCTCGGGGAGA	453		
QY	392	TCATCAGCATGGATGGGCTGCATCAGTGTGACCTTGGCAGCGCACCAAGGCTATTGGCTCA	451		
DB	454	TCATCAGCTGGATGCTCTATCAGGTGACCTTGGCAGCACACAGGCTATCGGCCTCA	513		
QY	452	AGGGGATCATTTGGCTGGCACTTGAGGAGCAGAAAGCCAAATACTTGCCTAACTGGCGT	511		
DB	514	AGGGGATTACTTTGGTCGGCAATCACGAACAGAAAGGCCAAGTATCTGCCCAAACATGTCGT	573		
QY	512	CCGGGAGCACATTTGAGCCTTCTGCTCAGGAGCCAGCGAGTGGGAGGATGACGCT	571		
DB	574	CTGGGGAACACATCCGAGCCTTCTGCCCTAACAGAGCCAGCGAGTGGGAGTGCATGCTGAT	633		
QY	572	CAATCCGGAGCAGAGCCACACTAAGTGAAGACAAAGACACTACATCTCAATGGCTCCA	631		
DB	634	CCATCCAGACGAGACTACGTTAAGTGAAGATAAGAGTACTTTCATCTCAATGGCTCCA	693		
QY	632	AGGTCTGGATTAATAATGGAGGACTGGCCAAATATTTTACTGTGTGGTGAAGACTGAGG	691		
DB	694	AGTCTGGATTAATAATGGAGGACTGGCCAAATATTTTACTGTGTGGTGAAGACTGAGG	753		
QY	692	TCGTTGATCTGATGGATCAGTGAAGAGCAAAATCAGAGCATTCATAGTAGAAGAGACT	751		
DB	754	TGGTTGATCTGATGGTTGCAAAACACAGCAAAATGACCGCATTCATAGTAGAAGAGACT	813		
QY	752	TTGGTGGAGTCACTAATGGGAAACCCGAGATAAATTAGGCATTTGGGGGCTCCAAACATT	811		
DB	814	TCGGCGGAATCACTAATGGGAAACCCGAGATAAATTAGGCATTCGAGGCTCCAAACAT	873		
QY	812	GTGAAGTCCATTTTGAACACCAAGATACCTGTGGAAAAACATCTCTCGAGAGGTCGGAG	871		
DB	874	GTGAGGTCCATTTTGAATAATACCAAGTGGCTGTGGAAAAATGTCCTCTGAGAGAGTTGAG	933		
QY	872	ATGGGTTTAAAGTGGCCATGAACATCTCTCAACAGCGCGCGGTTTCAGCATGGGAGCGTGC	931		

RESULT 12

BC031137

LOCUS

DEFINITION

BC031137 3071 bp mRNA linear ROD 10-JUN-2003
Mus musculus RIKEN cdna 2600017p15 gene, mRNA (cdna clone MGC:37218
IMAGE:4971848), complete cds.

Db	934	GGGGCTTTAAAGGTAGCCCATGAACATCTCTGAACAGTGGCGGATTCAGCATGGCAGTCTG	993
QY	932	TGGCTGGGCTGCTCAAGAGATTGTAAGATGACTGCTGAGTACGCCCTGCACAGGAAC	991
Db	994	TGGCCGGGATGCTGAAGAGCTGATCGAAGTACTGCTGAGTATGCTGCTGACGAGGAAC	1053
QY	992	AGTTTAAACAAGAGGCTCAGTGAATTTGGATTGATTCAGAGAAATTTGCATGATGGCTC	1051
Db	1054	AGTTCAACAGGAATCTCAGTGAATTTGGTTGATTTCAGGAAAAGTTTCGGCTCATGGCTC	1113
QY	1052	AGAAAGCTTACGTGAGAGATGACTACTCTACACAGCAGGAGTCTGGAACCAACCTG	1111
Db	1114	AGAAGGCATACGTAATGAGAGATAGGCTTACCTCCTTCAGGATGCTAGACCAACACAG	1173
QY	1112	GCTTTCCGACACTGCTCCATCGAGCAGCATGGTGAAGGTGTTTCAGCTCCGAGCGGCT	1171
Db	1174	GATTTCCGACACTGCTCTATCGAGCAGCATGGTGAAGGTGTTTACCTGAGGCTGCT	1233
QY	1172	GGCAGTGTGAGTGAAGGCTGAGATCTCGGGGCTTGGGCTACACAGGAGTACTC	1231
Db	1234	GGCAGTGTGAGCAGGCTCTGCAGATCTTTGGGGGCTCAGGCTACATGAAGGACTAC	1293
QY	1232	CGTAGCAGCATACTGGGTGACACCGCATCTCTCTATCTTCGAGGGAACCAATGAGA	1291
Db	1294	CTATGAGCGCATGCTGCGTGAATGCGCGCATCTCTCTATCTTCGAGGGAACCAATGAGA	1353
QY	1292	TTCTCCGGATGATACATCGCCCTGACGGGTCTGCAGATGCGCGGCTACCTCTGCTGCGG	1411
Db	1354	TTCTCCGGTGTTCATTTGCCCTGACAGGCTGCAGCATCTGCGAGCATCTCTGAGCTGCA	1413
QY	1352	GGATCATCAGCTTAAACAGCGCAAGTGCAGCAGCATGATGATACCGTTTGGCGGAGC	1411
Db	1414	GGATCAAGAGCTTAAAGTGGCAATGTGACCAAGTCAAGAGCAGTGTGCGGAAC	1473
QY	1412	TTCCGGACTCCCTGGGCGCAACTGTGGACTGGGCTTAACAGCGCATCTTGGTGTGTGCT	1471
Db	1474	TTCCGGATCTCTGGGCGCAACTGTGGACTGGGCTTAACAGCGCATCTTGGTGTGTGCT	1533
QY	1472	ACCCAGCTTTGGGAGCAGTGCACAAAGTTTGGAGAGAACACCTACTCTCTGCGCGGA	1531
Db	1534	ACCCAGCTTTGGGAGCAGTGCACAAAGTTTGGAGAGAACATGTCCTACTTTGGCGCGGA	1593
QY	1532	CGGTGAGACACTGCTGCTCCGCTTTGGCAAGACCATCATGAGGAGCAGTGGTACTGA	1591
Db	1594	CTGTTGAGACCTGTTACTCCGCTTTTGGAAAGAACATTTAGAGGAACAGCTGTTTAA	1653
QY	1592	AGCGGTGGCCAAACATCTCTCATCAACCTGTATGGCATGACGGCGTGTGTCGGGGCCA	1651
Db	1654	AGCGGTAGCCAAACATCTCTCATCAACCTGTATGGCATGACTGTGCTGTGAGAGCCA	1713
QY	1652	GGCGTTCATCCGCTTGGCTCCGCAACACAGCAGCAGGCTTCTTTGGCCCAACCT	1711
Db	1714	GGCGTTCATCAGAAATTTGGGCTCAGGAATCATGATCAGGAGTCTGTTGGCCCAATGT	1773
QY	1712	TCGCTGGAAGCTTACTTGCAGAAATCTCTTCAGCCTCTCTCAGCTGGACAAGTATGCTC	1771
Db	1774	TCGCTGGAAGCTTATTTCCAGATCTCTTCAGCCTGTCTCAGCTGGACAAGACGCTC	1833
QY	1772	CAGAAAACCTTAGATGAGCAGATTAAAGTGTCCAGCAGATCTCTTTGAGAACGCGCT	1831
Db	1834	CAGAAAACCTTAGATGAGCAGATTAAAGTGTCTCGCAGATCTCTTTGAGAACGCGCT	1893
QY	1832	ATATCTGTGCCACCTCTGAGCAGGACATGC	1863
Db	1894	ATATCTGTGCCACCTCTGAGCAGGACATGC	1925

ACCESSION

BC031137.1 GI:21410393

VERSION

MGC.

KEYWORDS

Mus musculus (house mouse)

SOURCE

ORGANISM

REFERENCE

AUTHORS

Eukaryote: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi, Muridae; Murinae; Mus. 1 (bases 1 to 3071)

Strausberg, R.D., Collins, F.S., Wagner, L.H., Derge, J.G., Klausner, R.L., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Altschul, S.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Hopkins, R.F., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uzdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mulligh, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, A., Whiting, M., Madan, A., Madan, A., Rodriguez, S., Bouffard, G.C., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2388257

12477932

2 (bases 1 to 3071)

Strausberg, R.

Direct Submission

Submitted (03-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 1IA03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabps@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>

Contact: (Dickson, Mark) mcdpaxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

FEATURES

source

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 59 Row: b Column: 9

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27369979.

Location/Qualifiers

1..3071

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:37218 IMAGE:4971848"

/tissue_type="Kidney, normal. 5 month old male mouse."

/clone_lib="NCI_CGAP_Kid14"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1..3071

/gene="2600017p15Rik"

/note="synonym: 4732402K02"

/db_xref="LocusID:229211"

/db_xref="MGI:1914272"

46..1923

gene

CDS

```

/codon_start=1
/product="2600017p15Rik protein"
/protein_id="AAH31137.1"
/db_xref="GI:21410394"
/db_xref="LocusID:229211"
/translacion="MSGCVLLSRGATAAAAAAARASRVLRFTARRRPLHTSLQSCSFA
KELFGNIRKQGVFPFPEVSHLSINOFVGLPKFTTFFEDVSRKIDQEGKIPVDTL
EKLSGLFGIQVPEEYGGLSNTMYARLGTIISLDASITVTLAAHOAIGUKGILV
GNDEOKAKYLPKLSGEHIAAFLCPETASGSDAASITQATISLSEDKKYFILNKSQWI
TNGGLANITFPAKTEVDSDSKDKTAPFIVERDFGINTGKEDKEDKLGIRGNTCE
VHEPTRVPEVNVLEGVGPKFVANNILNSMAYLTSGLMDQPPDCSIEAAMVVFSSSE
QNRNLSEFLQEPALQAFMAKQAYVNESMAYLTSGLMDQPPDCSIEAAMVVFSSSE
AAQCVCSEALQILGSGGYMKDYPYERMLRDARILLIFETNEILRLFTALGLQHAGR
ILTSRIKELSGNVTVMETIKRLRDLRTDGLGDLGVHPISLDSANKLEEN
VHYFGRTVETLLRFGKNIVEQLVKRVANILINLYGMTAVLSRASIRITGLRNHD
HBVLLANFCEVAYFQNLFLSLSOLDKAPENLDEQIKKVSRIKLEKRAYICAHPLDRA
S"
BASE COUNT      831 a      589 c      789 g      762 t
ORIGIN
Query Match      74.1%; Score 1380.8; DB 10; Length 3071;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 1550; Conservative 0; Mismatches 282; Indels 0; Gaps 0;
QY 32 CGGCTGCGGCTCGTCCGCTGCGGGGTCTGTTGGTCTCTACCCGCAACCGGGGCTACTGTC 91
Db 89 CGGCTGCGGCGCGGCTCTCCCGGGTCTGAGGGAATTCACGGCGAGGCGCGCGCTGC 148
QY 92 GCACGAGCGCGCTCTAGCAGCTTTTCGCCAAGAGCTTTCTCTAGGCAAAATCAAGAAGA 151
Db 149 ACACGAGCGCTGCAGAGCTGCTGTTCCGCCAAGAGCTCTTCTGCGGCAACATCAAGAGA 208
QY 152 AGAAGTTTTCCTCCATTTCCAGAGTTAGCCAGATGAACCTTAATGAATCAATCAGTTCT 211
Db 209 AAGGAGTTTTCCTCCATTTCCAGAGTTAGCCAGATGAACCTTAATGAATCAATCAGTTCT 268
QY 212 TGGGACCGCTGGAAAAATTTCTCACTGAAGAGTGGACTCCCGAAAAATTCACAGGAG 271
Db 269 TGGGACCGCTGGAAAAATTTCTCACTGAAGAGTGGACTCCCGAAAAATTCACAGGAG 328
QY 272 GAAATATCCAGATGAAATTTTGGAGAAATTTGAAGAGCTTAGGCTTTTGGGCTGCAAG 331
Db 329 GAAATATCCAGTTGACACCTTAGAGAGTTTGAAGAGCTTAGGCTTTTGGCATACAGG 388
QY 332 TCCAGAGATATATGTTGGCTGCGCTGCTCCACACACCATGACTCAAGACTAGGAGGA 391
Db 389 TCCAGAGATATATGTTGGCTGCGCTGCTCCACACACCATGACTCAAGACTAGGAGGA 448
QY 392 TCATCAGCATGGATGGGTCCATCATCTGTGACCTGGCAGCGCACAGGCTATTGGCCTCA 451
Db 449 TCATCAGCATGGATGGGTCCATCATCTGTGACCTGGCAGCGCACAGGCTATTGGCCTCA 508
QY 452 AGGGATCATCTTGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 511
Db 509 AGGGATCATCTTGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 568
QY 512 CCGGGAGCAGCATTCAGCTGAGCTTTCTGCTCAGCGAGCAGCAGTGGGAGCGTGCAGCCT 571
Db 569 CTGGGAGCAGCATTCAGCTGAGCTTTCTGCTCAGCGAGCAGCAGTGGGAGCGTGCAGC 628
QY 572 CAATCCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 631
Db 629 CCATCCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 688
QY 632 AGGTCTGATTAATGAGGAGCTGCGCAATATTTTACTGTGTTTGCAGAGCAGTGGAG 691
Db 689 AGGTCTGATTAATGAGGAGCTGCGCAATATTTTACTGTGTTTGCAGAGCAGTGGAG 748
QY 692 TCGTTGATTTCTGATGATCATGATGAAGAGCAAAATTCACAGCATTCATAGTAGAAGAG 751
Db 749 TGGTTGATTTCTGATGATTCGAAAGAGCAAAATTCACAGCATTCATAGTAGAAGAGAG 808
QY 752 TTGTTGAGTCACTAATGGGAAACCCGAGATAAATTTAGGAGTTCGGGGCTCCAGACTTT 811

```

Db	1889	ATATCTGTGCCACCCGCTGCAGAGCATCC	1920
RESULT 13			
BC001817			
LOCUS		1494 bp	mRNA
DEFINITION		1494 bp	linear
		IMAGE:3530094, mRNA, complete cds.	PRI 12-JUL-2001
ACCESSION		BC001817	
VERSION		BC001817.1	GI:12804754
KEYWORDS		MGC.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE		1 (bases 1 to 1494)	
AUTHORS		Straussberg, R.	
TITLE		Direct Submission	
JOURNAL		Submitted (29-JAN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK		NIH-MGC Project URL: http://mgc.nci.nih.gov	
COMMENT		Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbio.org contact: amadan@systemsbiology.org Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Kettelman and Anuradha Madan	
FEATURES		Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL plate: 11 Row: k Column: 19.	
source		Location/Qualifiers	
		1..1494	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="MGC:3368 IMAGE:3530094"	
		/tissue_type="Muscle, rhabdomyosarcoma"	
		/clone_lib="NIH MGC.17"	
		/lab_host="DH10B-R"	
		/note="Vector: pOTB7"	
		475..1041	
		/codon_start=1	
		/product="Similar to NPD002 protein"	
		/protein_id="AAH01817.1"	
		/db_xref="GI:12804755"	
		/translation="MYIALTGLQHAGRIITRIHELKQAKVSTVMDTVGRRRLDSILGR TVDLGTGNHGVVHPHSLADSKANKFEYTCFGRVETLLRFGKTIMEQLVLRVAN ILINFGMTAVLSRASRIRILNRHNDHLLANTPCVEAYLQNLFSLSQDKYAPEN LDBQIKKVSQILKRAYTCAHPLDRT"	
		370 a	
		386 c	
		422 g	
		316 t	
BASE COUNT			
ORIGIN			
		55.2%; Score 1029; DB 9; Length 1494;	
		Query Match	
		Best Local Similarity 100.0%; Pred. No. 6.7e-238;	
		Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	835	AAGATACCTCTGGAACACATCCTTGGAGAGTCGGAGATGGGTTTAAAGTGGCCATGAAC	894
Db	10	AAGATACCTCTGGAACACATCCTTGGAGAGTCGGAGATGGGTTTAAAGTGGCCATGAAC	69
QY	895	ATCCTCAACAGCGCCGGTTTACAGATGGGCGAGCGTCTGGCTGGCTCTCAGAGATTC	954
Db	70	ATCCTCAACAGCGCCGGTTTACAGATGGGCGAGCGTCTGGCTGGCTCTCAGAGATTC	129
QY	955	ATTCAATGACTCTGAGTACGCTGCACAGGAACAGTTTAAACAGAGGCTCAGTGA	1014

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2003, 10:14:13 : Search time 107.074 Seconds
(without alignments)
7679.711 Million cell updates/sec

Title: US-09-945-326-3

Perfect score: 1863

Sequence: 1 atgagcgcctcgggctctt.....accctctggacagacatgc 1863

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	168.6	9.0	1341	4	US-09-252-991A-11446
2	165.6	8.9	1212	4	US-09-252-991A-11389
3	155.6	8.4	1173	4	US-09-252-991A-9487
C 4	153.6	8.4	1209	4	US-09-252-991A-9346
C 5	143.4	7.7	4403765	3	US-09-103-840A-2
C 6	143.4	7.7	4411529	3	US-09-103-840A-1
C 7	139.4	7.5	3515	4	US-09-221-017B-824
8	134	7.2	1395	4	US-09-252-991A-16537
C 9	134	7.2	1665	4	US-09-252-991A-16061
10	132.6	7.1	33312	4	US-08-311-731A-121
C 11	132.2	6.6	1260	4	US-09-252-991A-14526
C 12	123.2	6.6	1260	4	US-09-252-991A-14838
13	111.2	6.0	1155	4	US-09-648-004-5
14	111.2	6.0	17417	4	US-09-648-004-27
15	105.8	5.7	1170	4	US-09-328-352-2254
16	102	5.5	1275	4	US-09-252-991A-8082
17	102	5.5	1488	4	US-09-252-991A-8036
C 18	102	5.5	2055	4	US-09-252-991A-8231
19	97	5.2	1185	4	US-09-328-352-2675
20	92.2	4.9	4403765	3	US-09-103-840A-1
21	92.2	4.9	4411529	3	US-09-103-840A-2
22	91.8	4.9	849	4	US-09-252-991A-12573
C 23	91.8	4.9	969	4	US-09-252-991A-12911
24	91.8	4.9	1257	4	US-09-252-991A-12881
25	90.4	4.9	1539	4	US-09-364-230-29
26	87.2	4.7	1164	4	US-09-328-352-2316
C 27	86.6	4.6	1440	4	US-09-252-991A-2675

28	86.6	4.6	2592	4	US-09-252-991A-3003	Sequence 3003, Ap
29	85.4	4.6	1553	4	US-09-364-230-31	Sequence 31, Appl
30	74.4	4.0	2079	4	US-09-252-991A-11426	Sequence 11426, A
31	74	4.0	1248	4	US-09-328-352-573	Sequence 573, App
32	72	3.9	1209	4	US-09-328-352-155	Sequence 155, App
33	70.8	3.8	1056	4	US-09-364-230-27	Sequence 27, Appl
34	70.2	3.8	1230	4	US-09-252-991A-14780	Sequence 14780, A
35	66.2	3.6	1182	4	US-09-252-991A-6534	Sequence 6534, Ap
C 36	66.2	3.6	1419	4	US-09-252-991A-6499	Sequence 6499, Ap
C 37	60	3.2	513	4	US-09-252-991A-9318	Sequence 9318, Ap
38	57.6	3.1	1149	4	US-09-328-352-724	Sequence 724, App
39	54.8	2.9	276	4	US-09-252-991A-9426	Sequence 9426, Ap
40	53.8	2.9	958	4	US-09-155-183-10	Sequence 10, Appl
41	53.8	2.9	1227	4	US-09-328-352-1449	Sequence 1449, Ap
42	53.4	2.9	766	4	US-09-364-230-33	Sequence 33, Appl
43	53.4	2.9	1200	4	US-09-328-352-234	Sequence 234, App
C 44	52.4	2.8	7218	1	US-08-232-463-14	Sequence 14, Appl
45	50.8	2.7	1302	4	US-09-252-991A-14660	Sequence 14660, A

ALIGNMENTS

RESULT 1

US-09-252-991A-11446/c

; Sequence 11446, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252.991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 11446

; LENGTH: 1341

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

; US-09-252-991A-11446

Query Match	9.08;	Score 168.6;	DB 4;	Length 1341;
Best Local Similarity	49.9%;	Pred. No. 5.2e-37;		
Matches 553;	Conservative 0;	Mismatches 529;	Indels 27;	Gaps 4;
Qy	242	AGGTGGACTCCGGAATAATTTGACCAAGGAGGAAATCCCGAGATGAACATTTGGGNAAT	301	
Db	1220	AGCGCGCGCGCGCAGATCGACGAGCAATCAGTTCCCATGGACATGTGGCGCANGT	1161	
Qy	302	TGAAGAGCCTTAGGGCTTTTGGCTGCAAGTCCCGAAGAAATATGTTGGCTGGGCTTCT	361	
Db	1160	TCGGCGAGATGGGCTGCTCGGCATCACCGTCGACGAGGAGTACGGCGGTTTCGGCGTGG	1101	
Qy	362	CCAAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGG---ATGGGTCCATCACTG	418	
Db	1100	GCTACCTGGCCCGCGCGCTGGTTCATGAGAGATACGCGGCGCTCGGCTCGGTGGCGC	1041	
Qy	419	TGACCTTGGCGAGCGCACAGGCTATTGGCCCTCAAGGGATCATCTTTGGCTGGCAGCTGAGG	478	
Db	1040	TCCTCTATGGCGGCGATTCACACTCTGCTCAACAGATCAAGCGCAACGCGCGC	981	
Qy	479	ACGAGAAAGCAATACTTGCCTAAACTGGCGTCGGGGAGCAGCATTTGACGCTTCTGCC	538	
Db	980	AACAGAGGCGCGCTACCTGCGCGCGCTGATATCCGCGCAACACATCGGTGCGCTGCCG	921	
Qy	539	TCAGGAGCGGCGGAGTGGGAGATGACGCTCAATCCGGAGCAGGACACTAAGTG	598	
Db	920	TGAGCGAACCACACGCGCGCTCGGAGCTGTCTCGATGAAGCTGCGCGCGGACCGGTGCG	861	

Db 3651995 GCGCCGAAACCGGTGAGCCGATCTGGGCTTCATTTGGCGGCTCGAAGTCTTCGCT 3651936
QY 1160 CCGAGGCGGCTGGCAGTGTGTGAGTGGCGCTGCAGATCCCTCGGGGCTTGGGCTACA 1219
Db 3651935 CCGAGTGGGATGGAGGTACACCGATGCCGTACAACCTGTTGGCGGCGCGGCTACA 3651876
QY 1220 CAAGGGCATATCCGTACGAGCGCATACTCGGTGACACCCGCGCTCTCTCTCATCTTCGAGG 1279
Db 3651875 CCACCGACTTCCCGGTGAGCGGTATATCGCGACGCCAAGATCACCCAGATCTATGAGG 3651816
QY 1280 GAACCAATGAGATCTCCCGATGTACATGCCCTGAGCGGTGCG 1324
Db 3651815 GCACCAATGAGATCTCCCGGTGAGTGTATGTCGCGGCGCTACTGC 3651771

RESULT 6

US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: TUBERCULOSIS
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 7.7%; Score 143.4; DB 3; Length 4411529;
Best Local Similarity 48.2%; Pred. No. 3.2e-28;
Matches 513; Conservative 0; Mismatches 531; Indels 21; Gaps 3;

QY 260 TTGACGAGAGGAAATCCAGATGAACCTTTGAGAAATGGAAGCCTAGGCGTTT 319
Db 3657964 TGGACGAGAGCGCGGTTCGCCGAAGAGCGCTAGTGGCGCTCAATTCCTCCGGTTTCA 3657905
QY 320 TTGGCTGCAAGTCCCGAAGAATATGTTGGCTGGCTTCTCCAAACACCATGTACTCAA 379
Db 3657904 ACGCGCTCCACATCCCGAGGAGTACGCGGTGAGGTCGCGTACTGCTACCTGCA 3657845
QY 380 GACTAGGGAGATCATCAGCATGGATGGTCCATCATCTGTGACCTGGCAGCGCACCGAG 439
Db 3657844 TCGTATCGAAGAGGTGGCGCGCTCGACGGTCTGCTGATCCCGCGGTCAACA 3657785
QY 440 CTATTGGCTCAAGGGATCATCTTGGCTGACCTGAGGAGCAGCAAGCAAAATCTGCG 499
Db 3657784 AGCTGGGACCATGGCTGTATCTGCGGGCTCGAGAGCTGAAGAAGCAGGTGCTGC 3657725
QY 500 CTAAGTGGCTCGGGGAGCACATTCAGCTTTCTGCTTCACGAGCGCACCGCTAGTGGGA 559
Db 3657724 CGGCGTTGGCGCGGAGGGCGATGGCTTCTATGATTCAGTGAACGCGCGCGGCA 3657665
QY 560 GCGATGAGCTCAATCCGGAGCAGACCATTAAGTGAAGACAGCAAGCATACATCC 619
Db 3657664 GTAGCGCGCTTCCATCGCGGACGCGGCGC-----AAGCGGATGGGATCATCTGGATTC 3657611
QY 620 TCAATGGCTCAAGGTCTGATTAATGAGGACTGGCAATATTTTACTGTGTTG 679
Db 3657610 TCAACGGCGCAAGTGTGATCATCAACGCGCAAGTCACTGTTACCGGTGATGG 3657551
QY 680 CAAAGCTAGGTGCTGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 739
Db 3657550 CGGTGACCGA-----TCCCGACCGGCGCGCAACGGCATCTCGCGGTTCATGG 3657503

QY 740 TAGAAGAGAGACTTTGTGTGAGTCACTAATATGGAACCCGAAAGATAAATTAGCATTCGGG 799
Db 3657502 TGCACAAGCAGCAGCAGGAGGTTCACCGTCCGTCGGAACCAACGCAAGCTCGGATCAAGG 3657443
QY 800 GCTCCAACACTTTGTGAAGTCCATTTTGAACACCAACAGATCTGTGGAACACATCTTTG 859
Db 3657442 GGTACCGACCAACCGAGCTGTACTTCGAGAATGCGCGATCCCGCGGATCGCATCATCG 3657383
QY 860 GAGAGTCGAGATGGTTTAAGTGGCCATCAACATCTCAACAGCGCGCGGTTTCAGCA 919
Db 3657382 GTGAGCCCGGTATCCGTTTCAAGACCGCGCTCGCCACGTTTGACACACACCCCTCCACGA 3657323
QY 920 TCGGAGCGCTGTGGCTGCTCAAGAGATTTGATGAATGATGCTGAGTACGCT 979
Db 3657322 TTGGCGCGCAGCGCGTGGGTATCGCCAGGCGCGCTGATGCTGCCATCGCTACACCA 3657263
QY 980 GCACAAGGAACAGTTTAACAAGAGGCTCAGTGAATTTGGATTTGATTCAGGAGAAATTTG 1039
Db 3657262 AGGACCGCAAGCAGTTTCGTGAGTGCATCAGTACGTTTTCAGGCCGTGAGTTCATGCTGG 3657203
QY 1040 CACTGATGCTCAGAAAGCTTACGTCATGGAGAGTATGACCTTACCTACAGCAGGATGC 1099
Db 3657202 CCGACATGCGATGAAGGTGGAGGCGCTCGGCTGATGTTTACTCCCGCGCGC---C 3657146
QY 1100 TGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGAGCGCATGGTGAAGTGTTCAGCT 1159
Db 3657145 GCGCGAAGCGGTGAGCGGATCGGCTTCATTTCCGCGCGCTTCGAAGTGTTCGCT 3657086
QY 1160 CCGAGCGCGCTGCGAGTGTGTGAGTGGCGCTGACAGATCTCGGGGCTTGGGCTACA 1219
Db 3657085 CCGAGTCCGATGAGGTACACCGATCGCTACAACTGTTGGCGCGCGCGGCTACA 3657026
QY 1220 CAAGGACTATCCGTACAGCGCATCTGCTGACACCCGCGATCTCTCTCTCTCTCTCTCT 1279
Db 3657025 CCACCGACTTCCCGTTCGAGCGGTTTATGTCGCGACGCCAAGATCACCGATCTATGAGG 3656966
QY 1280 GAACCAATGAGATCTCCGATGTACATCGCCCTGACGGGTCTGC 1324
Db 3656965 GCACCAATCAGATTCAGCGGCTAGTGTGTCGCGGCGCTACTGC 3656921

RESULT 7

US-09-221-017B-824/c
; Sequence 824, Application US/09221017B
; Patent No. 644799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIORITY DATA:
; APPLICATION NUMBER: PPI182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PPI546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:

Db	586	GTAGCAAGCACTTCAATCAGCCACCGCGGCACGCCGACTTCGCCATCGTTTCGCCGTCA	745
Qy	686	CTGAGGTCGTTGATTCTCTGATGGATCAGTGAAGAGCAAAATCACAGCAATTCATAGTAGAAA	745
Db	746	CCGACAGCTACGAGACAAACGCCGCAACGCGGTGACCGCTTCTCGTGGGACA	805
Qy	746	GAGACITTTGTTGGNGTCACTAATGGGNAACCCGAAGATAAATTAGGCAATTCGGGGGTCCA	805
Db	806	AGGGCACGCCCGGGATGACCGTGGCGCGCGGGCGGAAATCGGTGAGCAACCGTGGCTACC	865
Qy	806	ACACTTGTGAAGTCCAATTTTGAAGAACCAAGATACCTGTGGAAGAACATCCTTTGGAGAGG	865
Db	866	ACACCTACGAGATCTTCTTCAGCANVTGCCGGTGGCGGCTCCAAGGTGCTCGGGGAGG	925
Qy	866	TCGGAGATGGGTTTAAGTGGCCATGAACATCCTCAACAGCGGCCGTTTCAGCATGGCA	925
Db	926	TCGGCAAGGCTGGGAAGTGGCCAGCGCTGGCTCACCGCCGACGGGTGATGGTCGCCG	985
Qy	926	GCCTCGTGGCTGGGCTGCCTCAACAGATTGATTCAATGACTGCTGAGTAGCCCTGCACAA	985
Db	986	CCAACTGTGTCGGCCAGGCCCGCGGCTGACCTGTCGCTGGCGCGGCCGATC	1045
Qy	986	GGAACAGCTTTAAACAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCAC	1045
Db	1046	GCAAGCAGTTTCGCCAGCCGATCGGACGTACCAAGGCGTTTCTCTTCAAGCTGCCGACA	1105
Qy	1046	TGCTCAGAAAGCTTACGTCATCGAGATGATGACCTACCTCACAGCAGGAGTCTGGACC	1105
Db	1106	TGGCCACGAGATCCGTGGCGCGGCTGATGACCTGCACACCGCTGGAAGATGGACC	1165
Qy	1106	AACCTGGCTTTCGACACTGCTCCATCGAGCAGCCATGGTGAAGGTTTTCAGCTCCGAGG	1165
Db	1166	AGGGCACCATGACCGACGG-----CGAGCGCGCATGGCCAAGCTGTTGCCACGCGAGA	1219
Qy	1166	CCGCTTGGCAGTGTGTGAGTAGGCGCTCGAGATCTCGGGGGCTTGGGCTACACAAGG	1225
Db	1220	CCCTCGGCAGGTCGCGGACGAGCGGTCAGATCTTCGGCGGCATGGGCTGATGGATG	1279
Qy	1226	ACTATCCGTACGAGCGCATACTCGTGTACACCGCATCTCTCTCATCTTCGAGGGAACCA	1285
Db	1280	AAGGACCGTTCGAGCGCATCTGGCGCAACGCGGATCGAACGGATCTGGGAGGCACTT	1339
Qy	1286	ATGAGATTCCTCGGATGTACATCGCCCTGACGGGTCTGCAGC	1327
Db	1340	CGGAATCCAGCGGCACATCGTTTTCCTCCGGAACCTGCTCGGC	1381

RESULT 9

```

US-09-252-991A-16061/c
; Sequence 16061, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16061
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16061

```

Query Match 7.2%; Score 134; DB 4; Length 1665;
Best Local Similarity 48.8%; Pred. No. 2.4e-27;
Matches 430; Conservative 0; Mismatches 440; Indels 1

QY	446	GCCTCAAGGGGATCATCTTGCTGGCCTACTGAGGAGCAGAAGCAAATACTTGCCTAAAC	505
DB	1364	GGCCCTCGAAGATCCTCATGTCGCTGCACCGAGCAACAGCTCGCGACTAGCTGTGCCCCT	1305
QY	506	TGGCGTCGGGGAGCACATTCGACGCCTTCCTGCTACGAGCGCAGCCAGTGGGAGCGATG	565
DB	1304	CGGTGCAGGGGCAAAAGACCGATTGCTTCGCCCTTCACGAGCGCGGCGCGCTCGCAGC	1245
QY	566	CAGCCTCAATCCGGAGCAGGCCACACTAAGTGAAGCAAGAAGACACTATACATCCTCAATG	625
DB	1244	CCAATTCGATCAAGACCCGCGCGGTGCGCAGCGCGACGC-----CTTCGTGATCAACG	1191
QY	626	GCTCCAAGGTCTGATTACTAATGGAGGACTGSCCAATATTTTACTGTGTTTGCAGAGA	685
DB	1190	GTAGCAAGCACTTCATCAGCACCGCGGGCAGCCGACCTTCGGCATTGCTTCGCCGCTCA	1131
QY	686	CTGAGGTGCTTGATTCTTGATGGATCAGTGAAGACAAAAACACAGCAATTCATAGTAGAAA	745
DB	1130	CCGACAGCTACGAGCACACGCGCGCAAGCGCAACGCGGTGACCGCTTCCTGTGTGACA	1071
QY	746	GAGACTTTGGTGGAGTCACATAATGGGAACCCGAAGATAAATTTAGGCAATTCGGGGCTCA	805
DB	1070	AGGCAACCGCGGGATACCGCTGCGCGCGGGCGCGAATCGTGTAGCAACCGTGGCTACC	1011
QY	806	ACACTGTGGAAGTCCATTTTGAAGAACCAAGATACCTGTGGAAAAACATCTTTGGAGAG	865
DB	1010	ACACTACGAGATCTTCTCGACGATTGCGGGGTGCGGGCTCCAAGGTGCTCGCGCAGG	951
QY	866	TCGGAGATGGGTTTTAGGTGGCATGAACATCCTCAACAGCGCGCGTTTCAGCATGGCA	925
DB	950	TCGGCAAGGGCTGGGAATGGCCAAAGCCCTTGGCTCACCGCGGACGGGTGATGTCGCG	891
QY	926	CGCTGCTGGTGGCTGCTCAAGAGATTGATTCAAATGACTGCTGAGTAGCGCTGCACAA	985
DB	890	CCAACCTGTCTGGCCAGCCCGCGGCCCTCGACCTGTGCTGCGCTGGGCGGCCGATC	831
QY	986	GGAAACAGTTTAACAAGAGGCTCAGTGAATTTGGATTGATTCAGAGAAATTTGCACTGA	1045
DB	830	GCAAGCAGTTCGCCAGCCGATCGGCAGCTACCAAGGGCGTTTCTTCTCAAGCTCGCGCA	771
QY	1046	TGGCTCAGAGGCTTAGGTATGAGAGTATGACCTACCTCACAGCAGGGATGCTGGACC	1105
DB	770	TGGCCAGCAGATCCGTGCGCGCGAGCTGATGACCTGTGACACCGCGCTGGAAGATGGAC	711
QY	1106	AAGCTGGCTTTCCGAGCTGCTCCATCGAGCAGCCATGGTGAAGGTGTTTCAGCTCCGAGG	1165
DB	710	AGGCAACCATGACCGAGG-----CGAGCCGGCATGCCAAGCTGTTTCGCCAGCGAGA	657
QY	1166	CCGCTTGGCAGTGTGTGAGTGAGGCGCTCGAGATCCTCGGGGGCTTTGAGCGCTACACAAGG	1225
DB	656	CCCTCGGCAAGTCCGCGCAGAAACGGGTGCAGATCTTCGGCGGCATGGGCGCTGATGGATG	597
QY	1226	ACTATCCGTACGAGCGCATCTCGGTGACACCCGCATCCTCTCATCTTCAGGAGAACCA	1285
DB	596	AAGGACCGGTGAGCGCATCTGCGGCACACCGCGGATCGAACGGATCTCTGGGAGGCACTT	537
QY	1286	ATGAGATTCTCCGGATGTACATCGCCCTGACCGGCTGTCAGC	1327
DB	536	CGGAAATCCAGCGGCACATCTGTTTCCCGGGAACTGCTGCGGC	495

RESULT 10

```

US-08-311-731A-121
; Sequence 121, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: WOLF, GRENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,731A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: C0044/7125
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/720-3500
TELEFAX: 617/720-2441
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 33312 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: MYCOBACTERIUM LEPRAE
US-08-311-731A-121

Query Match	7.1%;	Score 132.6;	DB 4;	Length 33312;
Best Local Similarity	47.4%;	Pred. No. 2.7e-26;		
Matches 510;	Conservative	0;	Mismatches 544;	Indels 21; Gaps 3;
QY	260	TTGACCAGGAGGAAAATCCACGATGAAACTTTGGAGAAATTCGAAGAGCCTAGGGCTTT	319	
Db	24500	TGACACCGGTGCTGATTTCCGAGGAAGCGCTGGCAGCCCTGAATGCATCAGTTTCA	24559	
QY	320	TTGGGCTGCAAGTCCCGAAGAATATGTCGCTGGGCTTCTCCAAACACCAATGACTCAA	379	
Db	24560	ACGCTATCCACGTTCCCGAGGAGTATGGTGTGCTAGGCTGCGGATTCGGTAGTACGGCTTGCA	24619	
QY	380	GACTAGGGAGATCAATCAGCATGGATGGTCCATCACTGTGACCTCGCAGCGCACACAGG	439	
Db	24620	TTGTGATCGAAGAATGGCGGTGTGATGCTTCTGTGATCTGTGATTCCTGCAGTTTACA	24679	
QY	440	CTATTGGGCTCAAGGGGATCATCTTTGGCTGGCACTGAGGAGCAGAAAGCCAAATACTTGC	499	
Db	24680	AGCTTGGCACCATGGACTCATCTTCGCGGTTGGAAAGAGCTCAAGAAACAGGTTCTGC	24739	
QY	500	CTAAACTGGCTCCGGGAGACATATTGACGCTTCTGCCTCACGGAGCCACCGATGGGA	559	
Db	24740	CATCGTTGCTGCGGAGGGGCGATGGCGTCTCTATGCAATTAAGTGAAGCGCGAAGCCGGCA	24799	
QY	560	GCGATGCAGCCCTCAATCCGGAGCAGAGCCACACTAAAGTGAAGACAAGAAGCACATACATCC	619	
Db	24800	GTGACGCTGCTGATGATGGACCCCGGGCCAAA-----GCTGACGGGATGACTGGATTC	24853	
QY	620	TCAATGGCTCCAAAGTCTGGATTAATAATGAGGAGCTGGCCAAATATTTTACTGTGTTTG	679	
Db	24854	TCAATGGCTTCAAGTGCTGGATTACCAACGGTGGCAAGTCGACCTGGTACACAGGTTATGG	24913	
QY	680	CAAGACTCAGGTCTGTGATTCTGATGATCAGTGAAGACAAAATCACACATTCATAG	739	
Db	24914	CGGTGACCGATCCGACAAAGGGCGCA-----ACGGCATCTCGGCTTCATCG	24961	
QY	740	TAGAAGAGACTTTGGTGGAGTGCATAATGGGAAACCCGGAAGATAAATTAGGCATTCGGG	799	

```

RESULT 11
US-09-252-991A-14526
; Sequence 14526, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14526
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-14526

```

	Query Match	6.6%	Score 123.2	DB 4	Length 1260
	Best Local Similarity	47.8%	Pred. No. 2.1e-24		
	Matches 504	Conservative 0	Mismatches 523	Indels 27	Gaps 4
Qy	262	GACGAGGAAGGGAAAAATCCCAGATGAACATTTGGAGAAATTGAGAGCCCTAGGCGCTTTT	321		
Db	220	GACCGCGAGCACCGTTTTCCCGGGCGAGGCAATCGCGAGATGGCGCGACCTTGGCTTCCTC	279		
Qy	322	GGCGTCAAGTCCCAAGAAATATGTTGGCTGGGCTTCTCCAAACCAATGACTACAA--	379		
Db	280	GGATATCTGTTGCCAGGAATGGCGCGCGCGCGAGACCGGGCACCTTGGCTACGCCATG	339		

QY 1099 CTGGACCAACCTGGCTTTCCCGACTGCTCCATCAGGAGCCATGGTGAAGGTGTTCCAGC 1158
DB 225 CGGAA-----GCCGCTGCGGTGCTGACCGAGCCCTGATGCCAAGCTGTTCCGC 172
QY 1159 TCCGAGCGCCCTGGCAGTGTGTGAGTGAAGGCGCTGCAGATCCTCGGGGGCTTGGGCTAC 1218
DB 171 TCGGAGATGCCGAGGAAAGTTTGTCTCCGGCGGATCCAGACCTCGGTGGCTATGGCTAC 112
QY 1219 ACAAGGAGTATCCGTACGAGGCGCATCTGCTGACACCCGATCCTCTCATCTTCGAG 1278
DB 111 CTCAGGAGCTTCCCGTGAACGCATCTATCGGAGCGTGGGTCTGCCAGATCTACGAG 52
QY 1279 GGAACCAATGAGATTCCTCGGATGTACATCGGCC 1312
DB 51 GGCACCGAGCGCTGCGAGCGCTGGTGTGTCGCGC 18

RESULT 13
US-09-648-004-5
; Sequence 5, Application US/09648004
; Patent No. 6498242
; GENERAL INFORMATION:
; APPLICANT: CHEN, QIONG
; APPLICANT: THOMAS, STUART
; APPLICANT: NAGARAJAN, VASANTHA
; TITLE OF INVENTION: BIOLOGICAL METHOD FOR THE PRODUCTION OF ADIPIC ACID AND
; TITLE OF INVENTION: INTERMEDIATES
; FILE REFERENCE: CL-1341-A
; CURRENT APPLICATION NUMBER: US/09/648, 004
; CURRENT FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/252,553
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Acinetobacter sp.
US-09-648-004-5

Query Match 6.0%; Score 111.2; DB 4; Length 1155;
Best Local Similarity 47.8%; Pred. No. 4.5e-21;
Matches 499; Conservative 0; Mismatches 518; Indels 27; Gaps 5;

QY 275 AAATCCAGATGAACCTTTGGAGAAATTTGAAGAGCCTAGGCGCTTTTGGGCTGCAAGTCC 334
DB 104 AAATCCAGCTGAATCGTGCACCAATGAAGAACTGGGTCTTTTGGTCTCACCATTCC 163
QY 335 CAGAAGAATATGTGGCGCTGGCTTCTCCA---ACACCATGTACTCAAGACTAGGGAGA 391
DB 164 CTGAGGAATATGAGGGTCTTGGCTGACCATGGAGGAAGGTTTACATTGCAATTTGAAC 223
QY 392 TCATCAGCATGGATGGTCCATCATCTGACCTGCGAGCGCACAGGCTATTGGCTCA 451
DB 224 TGGGACGTACTCTCTGCTTCCGTTCACTGATCGCATAACATGGATGGATCGTTAT 283
QY 452 AGGGATCATCTTGGTGGCACTGAGGAGCAGAAACCAATATCTTGCCTAAACTGGCGT 511
DB 284 CAGGCTTAATATTGATGGCTCCGAGAGCAGAAACAGATATTTTGGCCACGCTGGCAA 343
QY 512 CCGGGAGCATATGAGCCTTCTGCTCAGCAGGCGCAGCAGTGGGAGGAGTGCACCT 571
DB 344 GTGGTGAATATTGGTTTCACTTCTGTTAACTGAACCTGATCCGGTTCAGATGCTGCC 403
QY 572 CAATCCGAGCAGCGCACTAAGTGAAGCAAGAGCACTACATCCTCAATGGGTCCA 631
DB 404 CTTT-----AAAACACAGCGGTGAAGATGGTGTATCATATTAAATGGCACTA 457
QY 632 AGGTCTGATTAATATGAGGACTGGCCAAATATTTTACTGTGTTTGCAGAACTGAGG 691
DB 458 AGCGTTACATCAACAAATGACCGCATCGGGTGTCTTTACTGTCTATGTCATGGCAGTAC----- 512

QY 692 TCGTTGATTTCTGATGGATCAGTGAAGACAAATCAGCAGCATTCATAGTAGAAAGAGACT 751
DB 513 ----CAGTACCAGAAATTAAGGTACAGGTGGATTTTACGCTTTTATCGTGGACAGTAAA 568
QY 752 TTGGTGGAGTCACTAATGGGAACCCGAGAGATAAATTAGGCATTCGGGCTCCACACATT 811
DB 569 CTCCTGGTATTTCCCTTGGGTAAACGTGATAAGAGATGGGCCAAAAGGTGCACATACCT 528
QY 812 GTGAAGTCCATTTTGAACACCAAGATACCTG---TGGAAAACATCTCTTGGAGAGGTGC 868
DB 629 GTGATGTGATTTTGAACACTTCGTATTCTCTGCATCTCACTCATTTGTTGTTGAAG 688
QY 869 GAGATGGGTTTAAAGTGGCCATGACATCCTCAACAGCGCGGTTTCACATGCGGACGCG 928
DB 689 GTGTAGGTTTAAACTGCAATGAGGTACTTTGATAAAGGCGGTATTCATATFTGCTGCAT 748
QY 929 TCGTGCCTGGCTGCTCAAGAGATGATTGAAATGACTGCTGAGTACGCTGCACAAGGA 988
DB 749 TAAGTGTAGTGTGCTACGCTGATGCTGGAAGATTCCTTACAATATGCGTTGAGCGCA 808
QY 989 AACAGTTTAAACAGAGGCTCAGTGAATTTGGATTGATTTCAGGAGAAATTTGCACCTGATGG 1048
DB 809 AACAGTTTGGTCAAGCGATTGCGAACTTCCAGTTGATTCAAGGTATGTTAGCCGATTTCTA 868
QY 1049 CTCAGAGGCTTACGTCATGAGAGATGATGACCTACCTCAGCAGGAGTCTGGACCAAC 1108
DB 869 AAGCTGAAATTTACGACAGCAAAATGATGTTATAGATGCTGCCACATTCGTGA----- 923
QY 1109 CTGGCTTTCCGACTGCTCCATCGAGGAGCCATGTTGAAGTGTTCAGCTCGGAGGCGG 1168
DB 924 -TGCTGGACAGATGTCAGCAGCGAAGCATCTTGTCCAGATGTTTGGCACTGAAATGT 982
QY 1169 CTTGGCAGTGTGTGAGTGAAGGCGCTGCAGATCCTCGGGGCTTGGGCTACAGAGGACT 1228
DB 983 GTGGCGCTGTCGAGATCGTGGCGTACAGATCCATGTTGGTGGGTTTATATCAGTGAAT 1042
QY 1229 ATCCGTACGAGCGCATACTGCTGACACCGCATCTCTCATCTTCGAGGGAACCAATG 1288
DB 1043 ATGCTATTGAGCGTTTTTACCGGTGATGTAGCTTTATTCGTTTGTATGAGGTACACGC 1102
QY 1289 AGATTTCTCGGATGTACATCGCC 1312
DB 1103 AAATCCACAGGTCATTATTGCC 1126

RESULT 14

US-09-648-004-27
; Sequence 27, Application US/09648004
; Patent No. 6498242
; GENERAL INFORMATION:
; APPLICANT: CHEN, QIONG
; APPLICANT: THOMAS, STUART
; APPLICANT: NAGARAJAN, VASANTHA
; TITLE OF INVENTION: BIOLOGICAL METHOD FOR THE PRODUCTION OF ADIPIC ACID AND
; TITLE OF INVENTION: INTERMEDIATES
; FILE REFERENCE: CL-1341-A
; CURRENT APPLICATION NUMBER: US/09/648, 004
; CURRENT FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/252,553
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 27
; LENGTH: 17417
; TYPE: DNA
; ORGANISM: Acinetobacter sp.
US-09-648-004-27

Query Match 6.0%; Score 111.2; DB 4; Length 17417;
Best Local Similarity 47.8%; Pred. No. 1.8e-20;
Matches 499; Conservative 0; Mismatches 518; Indels 27; Gaps 5;

QY 275 AAATCCAGATGAACCTTTGGAGAAATTTGAAGAGCCTTAGGGCTTTTGGGCTGCAAGTCC 334


```

Db 1880 AAATCCAGCTGAAATCTGCAGCAAAATGAAGAACTGGGTCTTTTGGTCTCACCAATC 1939
QY 335 CAGAAGAAATATGGTGGCTGGCTGGCTTCCCA---ACACATGTACTCAAGACTAGGGGAGA 391
Db 1940 CTGAGGAATATGAGGGTCTTGGCTGACCATGGAGGAAGAGTTTACATTTGCATTTGAAC 1999
QY 392 TCATCAGCATGGATGGGTCCATCACTGTACCCTGGCAGCGCACCAGGCTATTGGCCCTCA 451
Db 2000 TGGGACGTACCTCTCCTGCTTCCGTTCACTGATCGGCACATAACAATGGGATCGGTTTCAT 2059
QY 452 AGGGATCATCTTGGCTGGCACTGAGGAGCAGAAAGCCAAATACTTTCCTTAAACTGGCGT 511
Db 2060 CAGCTTAAATATTGATGGCTCCGAGAGCAGAAACAGATATTTTGGCACGCTGGCAA 2119
QY 512 CCGGGAGCAGCATTGCAGCCTTCTGCTCAGGAGCAGCAGCAGTGGGAGGATGACGCT 571
Db 2120 GTGGTGAATATTGGTTCAATCTGTTAACTGAACCTGATTCGGGTTCAGATGCTGCT 2179
QY 572 CAATCCGGAGCAGGCACACTAAGTGAAGACAGAAAGCACTACATCCTCAATGGCTCCA 631
Db 2180 CTTT-----AAAAACACAGCGTGAAGATGGTGATCATTTAATTTAATGGCACTA 2233
QY 632 AGTCTGGATTACTAATGGAGGACTGGCCAAATATTTTACTGTGTTTGAAGAGACTGAGG 691
Db 2234 AGCGTTACATCAACAATGCACCGCATCCGGGTCTTTTACTGTGTCATGGCAGCTAC----- 2288
QY 692 TCGTTGATTCTGATGATCAGTGAAGACAAAAATCACAGCATTCATAGTAGAAGAGACT 751
Db 2289 ----CAGTACCGAAATTAAGGTACAGGTGGAATTTTCAGCCTTTTATCGTGGAGCAAAA 2344
QY 752 TTGGTGGATCACTAATGGAAACCCGAAAGATAAATTAGGCAATTCGGGGCTCCACACAT 811
Db 2345 CTCCTGGTATTTCTTGGTAAACGTTGAAGAGATGGCCAAAGGTGCACATACCT 2404
QY 812 GTGAAGTCCATTTTGAAGAACACCAAGATACCTG---TGAAACATCTCTTGGAGAGTGG 868
Db 2405 GTGATGTGATTTTGAAGAACTGTGCTATTCCTGTCATCTGCATCTATGCTGTTGTTGAAG 2464
QY 869 GAGATGGTTTAAAGTGGCCATGAACATCCTCAACAGCGCGCGTTCAGCATGGGACGG 928
Db 2465 GTGTAGGTTTTAAACCTGCAATGAAGTACTTTGATAAGGCCGTATTCATATTTGCTGAT 2524
QY 929 TCGTGGCTGGCTGCTCAAGAGATTCATTTGAATGACTGCTGAGTACGCTGCACAAAGGA 988
Db 2525 TAAGTGTAGGTGCTGCTACGCGTATGCTGGAAGATTTCCCTACAAATGCGGTTGAGCGCA 2584
QY 989 AACAGTTTAAAGAGGCTCAGTGAATTTGGATTTGATTCAGGAGAAATTTGCACATGATGG 1048
Db 2585 AACAGTTTGGTCAAGCGATTGCGAACTTCCAGTTGATTCAGGTTATGATCCGATTTCTA 2644
QY 1049 CTCAGAGGCTTACGTGATGGAGATGATGACCTACCTCACAGCAGGGATGCTGGACCAAC 1108
Db 2645 AAGCTGAATTTACGACAGCAAAATGATGTTAGTATGATGCTGCCGCGACTTCGTGA----- 2699
QY 1109 CTGGCTTTCCGACTGCTCCATCGAGGAGCGCATGGTGAAGGTGTTACGCTCCGAGCGCG 1168
Db 2700 -TGCTGGACAGNATGACAGCGGAAGCATCTTGTGCCAAGATGTTTCCCACTGAAATGT 2758
QY 1169 CTGGCAGTGTGTAGTGAAGGCTGTCAGATCCTCGGGGGCTTGGGCTTACACAAAGGACT 1228
Db 2759 GTGGCGGTGTCGACAGATCCTGGCGTACAGATCCATGTTGGTGGCGGTTATATCATGTAAT 2818
QY 1229 ATCCGTACGAGCGCATCTGGTGCAGACCCCGCATCTCCTCATCTTCGAGGGAACCAATG 1288
Db 2819 ATGCTATTGAGCGTTTATACCGTGTATGATGATGATGATGATGATGATGATGATGATGATG 1312
QY 1289 AGATTCTCCGGATGATACATCCGCC 1312
Db 2879 AAATCCACACAGGTCAATATTGCC 2902

```

RESULT 15

```

US-09-328-352-2254
; Sequence 2254, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2254
; LENGTH: 1170
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-2254

```

```

Query Match 5.7%; Score 105.8; DB 4; Length 1170;
Best Local Similarity 47.3%; Pred. No. 1.4e-19;
Matches 502; Conservative 0; Mismatches 532; Indels 27; Gaps 5;
QY 257 AAATTCACGAGGAGGAAATCCAGATGAAACTTTGGAGAAATTTGAAGAGCCTAGGCG 316
Db 101 AAGTGGTAGAAGAACACAAATTCAGATGACATTTGTCAGCAATGCTGAGCTGGGTT 160
QY 317 TTTTGGGCTGCAAGTCCAGAAATATGTTGGCTGGCTTCTTCCAAACACCATGTACT 376
Db 161 TATTTGGCTCACAATTCGCAAGATATGTTGGCTTAGGCATCCTATGGAAGAGAG 220
QY 377 CAAGACT---AGGGAGATCATCAGCATGGATGGTGGTCCATCCTCTGTGACCTGGCAGCGC 433
Db 221 TCAGAGTTGCAATTTGAACTTGGACAACTTCACCGGCTTTTCGTTCTTTAATTTGGTACCA 280
QY 434 ACCAGGCTATTGGCTCAAGGGGATCATTTGGCTGGCACTGAGGAGCAGAGAAAGCCAAAT 493
Db 281 ATAACGGTATTGGCTCAAGTCAATTTTGGATGGCAGACAGAGCAAAAGCAGAAAT 340
QY 494 ACTTGCCTTAACTGGCTCCGGGAGCAGCATTTGAGGCTTCTGCTCAGGAGGAGCGCA 553
Db 341 ATTTACCAGCTTACGCAAGTGGCGAAATTTATGTTTCTATTTTAAACCGAGCTGAAT 400
QY 554 GTGGAGCGATGCAAGCCTCAATCCGGAGCAGGACACACATAGTGAAGACAGAGCAGCT 613
Db 401 CTGGTTCTGATGCTGCTCTTTTAAACACCATGCGCTA-----AAAGATGGCGATTTT 454
QY 614 ACATCCTCAATGGCTCCAGGCTCTGATGATTAATAGGAGCTGGCCAAATTTTACTG 673
Db 455 ATGATTAAACGGGCAACAAAGCTTTTATACCAATGCAACCCCATGCTGCAACTTTTACC 514
QY 674 TGTTCGAAAGACTGAGTCTGTTGATTTCTGATGATGATGATGATGATGATGATGATGATG 733
Db 515 TAATGGCTCGGAC-----GAATCCTGAAATTTAAGGGTCGGGTGGGATTTTCAGCTT 565
QY 734 TCATAGTAGAAGAGACTTTTGGTGGATCCTAATGGGAAACCCGAGATATAAATTAGGCA 793
Db 566 TTTTATTCGAGGCGCAATACACCGGCAATTTACATTTAGGCAAAATTTGACCAAAATGGG 625
QY 794 TTCGGGCTCCACACTTGTGAAGTCCATTTTGAACACCAAGATACCTG---TGGAAA 850
Db 626 AAAGGGTTCACATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 685
QY 851 ACATCCTTTGAGAGGCTCGGAGATGGGTTTAAAGTGGCAGTGAACATCTTCAACAGCGGCG 910
Db 686 TGATTGCTGCGTGAAGAGTGTGCGTTTAAACAGCGCATGAAAGTCTGGATGAAGGGC 745
QY 911 GGTTCAGCATGGCAGCGCTGCTGGCTGGCTGCTCAAGAGATTTGATTTGAATGACTGCTG 970
Db 746 GTTTGCACATTTGGTGGCTATAGCGTGGGTGTTGCCGAGCAGCATGTTTAAACGAGCAT 805
QY 971 AGTACGCTTCACAGGAAACAGTTTAAACAGAGGCTCAGTGAATTTGATTTGATTTGATTT 1030
Db 806 ACTATGGGATTGAGCGTTAAGCAGTTTGGTGCAGCCTATTGCGAACTTTTCAGCTTATTCAAG 865

```

Search completed: September 6, 2003, 12:58:37
Job time : 120.074 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2003, 23:10:50 ; Search time 3919.86 Seconds
(without alignments)
11551.252 Million cell updates/sec

Title: us-09-945-326-3

Perfect score: 1863

Sequence: 1 atgagcggctgcgggctctt.....accctctggacaggacatgc 1863

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1374.4	73.8	2432	11 AK075984	AK075984 Mus muscu
2	1371.8	73.6	2446	11 AK031820	AK031820 Mus muscu
3	1178.4	63.3	1914	11 BC041572	BC041572 Homo sapi
4	1074.6	57.7	2906	11 AK028571	AK028571 Mus muscu

5	937.8	50.3	1201	13 BX458691	BX458691
6	936.6	50.3	1071	12 BM808820	BM808820
7	910.2	48.9	1201	9 AL539220	AL539220
8	859.6	46.1	1010	12 BM552756	BM552756
9	845.4	45.4	1079	12 BM561207	BM561207
10	819.6	44.0	969	9 AL549834	AL549834
11	810.6	43.5	884	14 CA488487	CA488487
12	810	43.5	922	13 BU528453	BU528453
13	800.4	43.0	851	13 BU176038	BU176038
14	798.8	42.9	888	13 BQ960770	BQ960770
15	798.4	42.9	890	13 BQ938039	BQ938039
16	798	42.8	1201	9 AL560956	AL560956
17	797.6	42.8	937	13 BQ279135	BQ279135
18	797.4	42.8	1199	9 AL556077	AL556077
19	796.2	42.7	885	13 BU191988	BU191988
20	796.2	42.7	905	13 BU541780	BU541780
21	792.4	42.5	1023	13 BQ072180	BQ072180
22	780.2	41.9	891	13 BX328269	BX328269
23	777.2	41.7	914	9 AL524997	AL524997
24	772.4	41.5	851	10 BG757623	BG757623
25	772.2	41.4	875	13 BQ882994	BQ882994
26	772	41.4	1023	12 BM809053	BM809053
27	755	40.5	831	12 B1524078	B1524078
28	750	40.3	867	10 BG674212	BG674212
29	747	40.1	980	12 BQ054246	BQ054246
30	741.2	39.8	798	12 B1762564	B1762564
31	739.4	39.7	784	12 B1829609	B1829609
32	734.4	39.4	857	9 AL557535	AL557535
33	732.4	39.3	746	12 B1113771	B1113771
34	732.4	39.3	799	12 B1914477	B1914477
35	724.2	38.9	893	12 B1772602	B1772602
36	718.8	38.6	799	12 B1870667	B1870667
37	718.2	38.6	965	10 BG109786	BG109786
38	717.2	38.5	973	10 BE792118	BE792118
39	717	38.5	845	13 BU153157	BU153157
40	716	38.4	748	12 BG764035	BG764035
41	707	37.9	913	10 BE795758	BE795758
42	704.6	37.8	771	9 AU142885	AU142885
43	704	37.8	745	12 BG760511	BG760511
44	703.6	37.8	942	13 BX350634	BX350634
45	702.8	37.7	832	12 B1911835	B1911835

ALIGNMENTS

RESULT 1
AK075984
LOCUS
DEFINITION
AK075984. 2432 bp mRNA linear HTC 07-DEC-2002
Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:2600017P15 product:VERY-LONG-CHAIN ACYL-COA
DEHYDROGENASE VLCAD homolog [Homo sapiens], full insert sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AK075984.1 GI:26344893
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, N., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED
11042159

Db	552	CTGGGGAACACATCCGACCTTCTCCTAACAGACGACGCGAGTGGAGTGATGCTCGT	611
Qy	572	CAATCCGAGCAGACACACTAAGTGAAGACAAGACCACTACATCCTCAATGGCTCCA	631
Db	612	CCATCCAGCAGACACTAGTTAAGTGAAGATAAGAAGTACTTCATACTCAATGGCTCCA	671
Qy	632	AGGTCTGGATTACTAATGGAGGACTGCCCAATATTTTACTGTGTTTGCAAAAGACTGAGG	691
Db	672	AGSTTTGGATCACCAGTGGAGGACTGCCCAATATTTTACTGTGTTTGCAAAACCCAGG	731
Qy	692	TCGTTGATCTGATGGATCAGTGAACACAAAATCAGCATTCATAGTAGAAGAGACT	751
Db	732	TGTCGATCTGATGGTTCGAACACAGACAAAATGACCGCATTCATAGTAGAAGAGACT	791
Qy	752	TTGGTGGAGTCACTAATGGGAACCCGAAGATAAATTAGGCATTCGGGGCTCCAAACACTT	811
Db	792	TCGGCGGAATCACTAATGGGAACCTGAAGATAAATTAGGCATTCGAGGCTCCACACAT	851
Qy	812	GTGAAGTCCATTTTGAACACCAAGATACCTGTGGAACACATCCTTTGGAGAGGTCGGAG	871
Db	852	GTGAGGTCCATTTTGAACATACAGAGTGCCTGTGGAATATGTCTTGGAGAAGTTGGAG	911
Qy	872	ATGGTTTAAGGTGGCCATCAACATCTCAACAGCGCGGTTTCAGCATGGGAGCGTCG	931
Db	912	GGGCTTAAAGTAGCCATGAACATCTGAACAGTGGGCAATTCAGCATGGGAGCTGCTG	971
Qy	932	TGGCTGGCTGCTCAAGAGATTAATTTGAATGACTGCTGAGTACGCTGCACAAAGAAAC	991
Db	972	TGGCGGGATGCTGAAGAGCTGATCGAATGACTGCTGAGTATGCTGTACGAGAAAC	1031
Qy	992	AGTTTAAACAGAGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCACTGATGGCTC	1051
Db	1032	AGTTCAACAGGAATCTCAGTGAATTTGGTTGATCCAGGAAGATTTGGCTCATGGCTC	1091
Qy	1052	AGAAGCTTACGTCATGGAGATGATGACCTACCTCACAGGAGTCTGGACCACTG	1111
Db	1092	AGAAGCATACGTAATGGAGAGTATGCTTACCTCACTTCAGGATGCTAGACCAACAG	1151
Qy	1112	GCTTCCGACGCTCCATCGAGGAGCCATGGTGAAGTGTTCAGTCCGAGGCGCT	1171
Db	1152	GATTTCCGACGCTCTATCGAGGAGCCATGGTGAAGTGTTCAGTCCGAGGCGCT	1211
Qy	1172	GGCAGTGTGTGAGTGGGCGCTGCAGATCCTCGGGGCTTGGCTACACAGGAGTATC	1231
Db	1212	GGCAGTGTGTGAGGAGGCTTCAGATCCTTTGGGGCTTCAGGCTACATGAAGAGTACC	1271
Qy	1232	CGTACAGGCGCATCTGCTGCACACCCGATCCTCTCATCTTCAGGGAACCAATGAGA	1291
Db	1272	CCTATGAGGCGCATCTGCTGCATGATGCCGATCTCTCTATCTTCAGGGAACCAATGAGA	1331
Qy	1292	TTCTCCGATGTACATCGCCCTGACGGTCTGCAGCATCCCGGCGCATCTCTGACTACA	1351
Db	1332	TTCTTCCGTTGTTCTATGTCCTTCAGAGGCTTCAGCATCTGCAGGCGCATCTCTGACTCGA	1391
Qy	1352	GGATCCATGAGCTTAAACAGGCCAAGTGCAGCATGATGATACCTGTGGCGGAGGC	1411
Db	1392	GGATCAAGAGCTTAAAGTGGCAATGTGACCACTCATCGAGACGATTTGGTCGGAAC	1451
Qy	1412	TTCCGGACTCCTTGGCGCACTGTGGACCTTGGGCTGCAGGCAACCATGGATGTGC	1471
Db	1452	TTCCGGATCTCTTGGCGCACTGTGGACCTTGGGCTTAACAGGCGATCTTGGTGTGTC	1511
Qy	1472	ACCCCGACTCTTCCGACAGTGCACAAAGTTTGGAGGACCAACCTACTGCTTCGGCGGA	1531
Db	1512	ACCCCGACTTGGCCAGAGTGCACAAAGCTTGGAGGAAATGTCCATTTTGGCGGA	1571
Qy	1532	CGTGGAGACTGCTGCTCCGCTTTGGCAAGACCATCATGGAGGAGCAGCTGTACTGA	1591
Db	1572	CTGTTGACCCCTGTTACTCCGTTTGGAAAGAACATTTAGAGGAACAGCTGTGTAA	1631
Qy	1592	AGCGGGTGGCAACATCTCATCACTGATGGATGACGGCGCTGTCTGCGGCGCA	1651
Db	1632	AGCGGGTAGCCAAACATCTCATCACTTGTATGTCATGCTGCTGTCTGTCGAGAGCCA	1691
Qy	1652	GCGCTCCATCCGATTTGGCTCCGCAACACGACGAGGTTCTTCTTTGGCCACACCT	1711
Db	1692	GCGCTCCATCCGATTTGGCTCCGCAACATCATGATCAGGAGTCTTCTTTGGCCACACCT	1751
Qy	1712	TCGCTCGGAGGCTTACTTTGCGAATCTCTTACGCTCTCTCAGCTGGACAGTAGTGTCTC	1771
Db	1752	TCGCTCGGAGGCTTACTTTCCAGAACTCTTACGCTCTCTCAGCTGGACAGTAGTGTCTC	1811
Qy	1772	CAGAAACCTAGATGAGCAGATTAGAAAGTGTCCGAGCAGATCCTTTGAGAGCGAGCT	1831
Db	1812	CAGAAACCTGATGAGCAGATTAGAAAGTGTCTCGGAGATCCTTTGAGAGCGAGCT	1871
Qy	1832	ATATCTGTCGCCACCTCTCGGACAGCAGATGC	1863
Db	1872	ATATCTGTCGCCACCTCTCGGACAGCAGATGC	1903
RESULT 2			
AK031820			
LOCUS			
DEFINITION			
AK031820		2446 bp	mRNA
LINEAR			linear
HTC			HTC 05-DEC-2002
ENRICHED			musculus adult male medulla oblongata cDNA, RIKEN full-length
LIBRARY			enriched library, clone:6330407H09 product:VERY-LONG-CHAIN ACYL-COA
DERIVATION			DERIVATION: VLDLAD homolog [Homo sapiens], full insert sequence.
ACCESSION			AK031820
VERSION			AK031820.1 GI:26327642
KEYWORDS			HTC; CAP trapper.
SOURCE			Mus musculus (house mouse)
ORGANISM			Mus musculus
REFERENCE			1
AUTHORS			Carninci, P. and Hayashizaki, Y.
TITLE			High-efficiency full-length cDNA cloning
JOURNAL			Meth. Enzymol. 303, 19-44 (1999)
MEDLINE			99279253
PUBMED			10349636
REFERENCE			2
AUTHORS			Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE			Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL			Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE			20499374
PUBMED			11042159
REFERENCE			3
AUTHORS			Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
TITLE			RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipeptide sequencing
JOURNAL			Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE			20530913
PUBMED			11076861
REFERENCE			4
AUTHORS			Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schirml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Anon, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,

433 TGGCTGGGCTGCTCAAGAGATTGATTGGTAGAAATGACTGCTGAGTACGCTGCACAAAGG 492
988 AAACAGTTTAAACAGAGGCTCAGTGAATTTGATTGATT----- 1026
493 AAACAGTTTAAACAGAGGCTCAGTGAATTTGATTGATTGATTCTTTCTTTTCTG 552
1027 -----CAGGAGAAATTTGCACTGATCGCTCAGAAAGGCTTACGTCATGGAGATGATGAC 1079
553 AACACTCCAGGAGAAATTTGCACTGATCGCTCAGAAAGGCTTACGTCATGGAGATGATGAC 612
1080 CTACCTCAGCAGGAGATGCTGGACCAACCTGGCTTTCCCGACTGCTCATCGAGGAGC 1139
613 CTACCTCAGCAGGAGATGCTGGACCAACCTGGCTTTCCCGACTGCTCATCGAGGAGC 672
1140 CATGTTGAAGTTTCAGTCCGAGCGCCCTGGCGAGTGTGAGTGGAGGCTCGAGT 1199
673 CATGTTGAAGTTTCAGTCCGAGCGCCCTGGCGAGTGTGAGTGGAGGCTCGAGT 732
1200 CTTGCGGGGCTTGGGCTACACAGGAGCTATCCGTACGAGCGCATCTGCGTGCACCCG 1259
733 CTTGCGGGGCTTGGGCTACACAGGAGCTATCCGTACGAGCGCATCTGCGTGCACCCG 792
1260 CATCTCTCTATCTTCGAGGGAACCAATGAGATTCTCCGATGTACATCGCCCTGACGG 1319
793 CATCTCTCTATCTTCGAGGGAACCAATGAGATTCTCCGATGTACATCGCCCTGACGG 852
1320 TCTGCAGCATGCGGCGCATCTGACTACAGGATCCATGAGCTTAAACAGGCCAAAGT 1379
853 TCTGCAGCATGCGGCGCATCTGACTACAGGATCCATGAGCTTAAACAGGCCAAAGT 912
1380 GAGCAGCATGATGATACCTGTTGGCGGAGGCTTGGGACTCCCTGGGCGCACTGTGA 1439
913 GAGCAGCATGATGATACCTGTTGGCGGAGGCTTGGGACTCCCTGGGCGCACTGTGA 972
1440 CTTGGGGCTGACAGGCAACCATGGAGTTGTGCACCCAGCTTGTGGCAGATGCCAACAA 1499
973 CTTGGGGCTGACAGGCAACCATGGAGTTGTGCACCCAGCTTGTGGCAGATGCCAACAA 1032
1500 GTTTGAGGAGAACACCTACTGCTTTCGCGCGGACCGTGGAGACACTGCTGCTCGCTTGG 1559
1033 GTTTGAGGAGAACACCTACTGCTTTCGCGCGGACCGTGGAGACACTGCTGCTCGCTTGG 1092
1560 CAAGACCATGAGGAGCAGTGTACTGAGCGGGTGGCCAACTCTCATCAACCT 1619
1093 CAAGACCATGAGGAGCAGTGTACTGAGCGGGTGGCCAACTCTCATCAACCT 1152
1620 GTATGGCATGACGGCGTGTGTCGCGGCGACCGCTCCATCCGATTTGGCTCGCA 1679
1153 GTATGGCATGACGGCGTGTGTCGCGGCGACCGCTCCATCCGATTTGGCTCGCA 1212
1680 CCAGCAGCAGGATTTCTTTGCGCAACACCTTCTCGGTGGAGCTTACTTGCAGATCT 1739
1213 CCAGCAGCAGGATTTCTTTGCGCAACACCTTCTCGGTGGAGCTTACTTGCAGATCT 1272
1740 CTTGAGCTCTTCAGTGCAGAGATGCTCCAGAAAACCTAGATGAGCAGATTAAGAA 1799
1273 CTTGAGCTCTTCAGTGCAGAGATGCTCCAGAAAACCTAGATGAGCAGATTAAGAA 1332
1800 AGTGTCCAGCAGATCTTTCAGAGGAGGAGCTTATCTGTGCCACCCCTTGTGCAGGAG 1859
1333 AGTGTCCAGCAGATCTTTCAGAGGAGGAGCTTATCTGTGCCACCCCTTGTGCAGGAG 1392
1860 ATGC 1863
1393 ATGC 1396

RESULT 4
AK028571
LOCUS
DEFINITION
Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:4732402K02 product:VERY-LONG-CHAIN ACYL-CoA DEHYDROGENASE VLCD homolog [Homo sapiens], full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

AK028571
AK028571.1 GI:26080919
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yananaka, I., Saito, K., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nika, I., Nika, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamei, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Sakamoto, N., Sasaki, H., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weltz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaka, S., and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

6 (bases 1 to 2906)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imetani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, K., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,

TITLE	COMMENT
<p>JOURNAL</p> <p>Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Taya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.</p> <p>Direct Submission</p> <p>Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Fax: 81-45-503-9216)</p> <p>cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.</p> <p>Please visit our web site for further details.</p>	

[illegible]

QY	927	CGTCTGGCTGGCGCTGCTCAAGAGATTTGATTGAATGACTGCTGAGTACGCTGCACAAG	986
Db	1762	TGCTGTGGCGGGATGCTGAAGAGCTGATGAACTGACTGCTGAGTATGCCCTGACGAG	1821
QY	987	GAACAGTTTAAACAAGAGGCTCAGTGAATTTGGATTGATTCAGAGAAATTTGCACATGAT	1046
Db	1822	GAACAGTTCAACAGGAATCTCAGTGAATTTGGTTGATCCAGGAAAGTTTGGCGTCAT	1881
QY	1047	GGCTCAGAAGGTTTACGTCATGGAGAGTATGACCTTACCTCACAGCAGGGATGCTGGACCA	1106
Db	1882	GGCTCAGAAGGCATACGTAATGGAGAGTATGGCTTACCTCACTTCAGCGGATGCTAGACCA	1941
QY	1107	ACCTGGCTTTCGGACTGCTCCATCGAGGAGCGCCATGGTGAAGGTGTTGAGCTCCGAGGC	1166
Db	1942	ACCAGATTTCCGACTGCTCTATCGAGGAGCGCCATGGTGAAGGTGTTTAACTCTGAGGC	2001
QY	1167	CGCCTGGCAGTGTGTGATGAGGGCTGCGAGATCCTCGGGGGCTTGGGCTTACACAAGGA	1226
Db	2002	TGCCCTGGCAGTGTGTGAGCGAGGCTCTGCAGATCCTTGGGGCTCAGGCTACATGAAGGA	2061
QY	1227	CTATCCGTACAGCGCATCTGCGTGACACCCGCATCCTCCTCATCTTCAGGGAACCAA	1286
Db	2062	CTACCCCTATGAGCGCATGCTGCGGTGATGCCGCATTTCTCTAATCTTTGAGGGAACCAA	2121
QY	1287	TGAGATTCTCCGGATGTACATCGCCCTGAGGGGTCTGCAGCATCGCGGCGCATCTCTGAC	1346
Db	2122	TGAGATTCTTCGGTTGTTCAATGCGCTGCAGGCTCCAGCATGCTGGAGCATCCTGAC	2181
QY	1347	TACCAGATCCATCAGCTTAAACAGGCCAAAGTGGAGACAGTCAATGATACCGTTGGCCG	1406
Db	2182	CTCAGAGATCAAGAGCTTAAAGTGCGAATGTGACCCAGTCAATGAGACGATTTGGTCG	2241
QY	1407	GAGCTTTCCGACTTCCCTGGCGCGGAAGTGTGGACCTGGGGCTGCACAGCAACCATGGAGCT	1466
Db	2242	GAACCTTCGGGATTTCTCTGGCGCGAAGTGTGGACCTGGGGCTAACAGGCGCATCTGGTGT	2301
QY	1467	TGTGCACCCAGTCTTCGGACACAGTGCACAACAGTTTGAAGGAGAACCATCTGCTTCGG	1526
Db	2302	TGTCCACCCAGGCTTGGCCACAGTGCACAACAGCTTGAAGGAAATGTCCATTACTTTGG	2361
QY	1527	CCGACCGTGGAGACACTGCTGCTCCGCTTTGGCAAGACCATCATGAGGAGCACGCTGT	1586
Db	2362	CCGACCTGTTGAGACCTGTTACTCCGTTTGGAAAGAACATTTAGAGGAACACGCTGGT	2421
QY	1587	ACTGAAGCGGGTGGCCAAACATCTCATCAACCTGTATGGCATGAGCGCGTCTGCTCGCG	1646
Db	2422	GTTAAGCGGGTAGCCAAACATCTCATCAACTTGTATGGCATGACTGTGCTGCTCGAG	2481
QY	1647	GGCAGCGCTCCATCCGATTGGGCTCCGCAACACGACACGAGGTTCCTCTGGCCAA	1706
Db	2482	AGCCAGCGCTCCATCAGAATTTGGCTCAGGAATCATGATCACGAGGTCTGTTGGCCAA	2541
QY	1707	CACCTTCTCGGTGGAGCTTACTTGCAGAAATCTTTCAGCTTCCTCAGCTGGACAAGTA	1766
Db	2542	CATGTTCTGTGTGGAAAGCTTATTTCCAGAAATCTCTTCAGCCTGTCTCAGCTGGACAAGAA	2601
QY	1767	TGCTCCAGAAAACCTAGATGAGCAGATTAAAGAAAGTGTCCCAAGCAGATCTCTTGAGAGCG	1826
Db	2602	CGCTCCAGAAAACCTGGATGAGCAGATTAAAGAAAGTGTCTCGGAGATCTTTGAGAGCG	2661
QY	1827	AGCCTATATCTGTGCCACCCCTCTTGACAGGACATGC	1863
Db	2662	AGCCTATATCTGTGCCACCCCTTGACAGGACATCC	2698
RESULT	5		
LOCUS	1201	bp	
DEFINITION	linear	mRNA	EST 22-MAY-2003
ACCESION	1		
VERSION	1		
GI	31025045		

KEYWORDS SOURCE ORGANISM	EST.		Homo sapiens (human)	
	Homo sapiens		Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	1 (bases 1 to 1201)		Li, W.B., Gruber, C., Jessee, J., and Polayes, D.	
REFERENCE AUTHORS TITLE JOURNAL COMMENT	Full-length cDNA libraries and normalization		Unpublished	
	Contact: Genoscope		Genoscope - Centre National de Sequencage	
	BP 191 91006 EVRY cedex - France		Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr	
	Library was constructed by Life Technologies, a division of		Invitrogen. This sequence belongs to sequence cluster 2225.f For	
FEATURES source	more information about this cluster, see		http://www.genoscope.cns.fr/	
	cgi-bin/cluster.cgi?seq=CS0DE007AF12QP1&cluster=2225.f. Contact :		Feng Liang Email : fliang@lifetech.com URL :	
	http://fulllength.invitrogen.com/ Invitrogen Corporation 1600		Paradise Avenue Genoscope sequence ID : CS0DE007AF12QP1.	
	Location/Qualifiers		1. .1201	
BASE COUNT ORIGIN	290 a 284 c 337 g 240 t		50 others	
	Query Match		50.3%; Score 937.8; DB 13; Length 1201;	
	Best Local Similarity		97.6%; Pred. No. 3.5e-246;	
	Matches 983; Conservative 9; Mismatches 11; Indels 4; Gaps 4;			
QY	712 GTGAAGACAAATACAGACATCATAGTAGAAGACATTTGTGTGAGTCACTAATGGG 771			
	DB		66 GTGAAGACAAATACAGACATCATAGTAGAAGACATTTGTGTGAGTCACTAATGG 125	
	QY		772 AAACCCGAAGATAAATAGGCATTTCGGGGCTCCAACTGTGAAGTCCATTGGAAC 831	
	DB		126 AAACCCGAAGATAAATAGGCATTTCGGGGCTCCAACTGTGAAGTCCATTGGAAC 185	
QY	832 ACCAAGATACCTGTGGAACATCCTTGGAGAGTCGGAGATGGGTTTAAAGTGGCCATG 891			
	DB		186 ACCAAGATACCTGTGGAACATCCTTGGAGAGTCGGAGATGGGTTTAAAGTGGCCATG 245	
	QY		892 AACATCTCTCAACAGCGCGCTTCAGATGGGCGAGCGCTGCTGGCTGGCGCTCAAGAGA 951	
	DB		246 AACATCTCTCAACAGCGCGCTTCAGATGGGCGAGCGCTGCTGGCTGGCGCTCAAGAGA 305	
QY	952 TTGATTGAATGACTGTGAGTACGCTGTCACAGGAACAGTTTACACAGCGCTCAGT 1011			
	DB		306 TTGATTGAATGACTGTGAGTACGCTGTCACAGGAACAGTTTACACAGCGCTCAGT 365	
	QY		1012 GAATTGTGATTGATTTCAGAGAAATTTGCATGTGCTCAGAGAGGCTTACGTCATGGAG 1071	
	DB		366 GAATTGTGATTGATTTCAGAGAAATTTGCATGTGCTCAGAGAGGCTTACGTCATGGAG 425	
QY	1072 AGTATGACCTACCTACAGCAGGGATGCTGACCAACCTGGCTTCCCGACTGCTCCATC 1131			
	DB		426 AGTATGACCTACCTACAGCAGGGATGCTGACCAACCTGGCTTCCCGACTGCTCCATC 485	
	QY		1132 GAGGACGACCTGTTGAAGTGTTCAGTCCGAGCGCGCTGCGAGTGTGTCAGTGAGGCG 1191	
	DB		486 GAGGACGACCTGTTGAAGTGTTCAGTCCGAGCGCGCTGCGAGTGTGTCAGTGAGGCG 545	
QY	1192 CTGCAGATCCTCGGGGGCTTGGGCTACACAGGAGACTATCGTAGGAGCGATACGCT 1251			
	DB		546 CTGCAGATCCTCGGGGGCTTGGGCTACACAGGAGACTATCGTAGGAGCGATACGCT 605	
	QY		1252 GACACCCGACCTCCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCC 1311	
	DB		606 GACACCCGACCTCCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCC 665	
QY	1312 CTGACGGGTCTGCAGCATGC-CGGCCGCGATCCTGACTACCAAGGATCCATGAGCTTAAACA 1370			
	DB		666 CTGACGGGTCTGCAGCATGC-CGGCCGCGATCCTGACTACCAAGGATCCATGAGCTTAAACA 725	
	QY		1371 GCCCAAGTGCAGCAGTATGATACCTGTCGGCGAGGCTTCGGGACTCCCTGGGCGG 1430	
	DB		726 GCCCAAGTGCAGCAGTATGATACCTGTCGGCGAGGCTTCGGGACTCCCTGGGCGG 784	
QY	1431 AACTGTGGACCTGGGGCTGCACAGCAACCATGAGTGTGACCCAGTCTTTCGGGAGAG 1490			
	DB		785 AACTGTGGACCTGGGGCTGCACAGCAACCATGAGTGTGACCCAGTCTTTCGGGAGAG 844	
	QY		1491 TGCCCAAGTGTGAGGAGAACCTACTGCTTCGGCGGAGCGTGGAGACACTGCTGCT 1550	
	DB		845 TGCCCAAGTGTGAGGAGAACCTACTGCTTCGGCGGAGCGTGGAGACACTGCTGCT 904	
QY	1551 CCGCTTGGCAAGACCATCATGAGGAGCAGCTGTGACTGAAAGCGGGTGGCCAAACATCCT 1610			
	DB		905 CCGCTTGGCAAGACCATCATGAGGAGCAGCTGTGACTGAAAGCGGGTGGCCAAACATCCT 964	
	QY		1611 CATCAACCTGTATGGCATGACGGCGCTGCTGCGGGGCGAGCGCTCCATCCGATGGG 1670	
	DB		965 CATCAACCTGTATGGCATGACGGCGCTGCTGCGGGGCGAGCGCTCCATCCGATGGG 1024	
QY	1671 GCTCCGCAACCGACGACGAGGTTCTCTTGGCCAAACACTTCTCGG 1717			
	DB		1025 GST-CGCAACMACGA-CACGAGGKTTTGTGGCCAAACCTTTTSGTG 1069	
RESULT 6 BM808820 LOCUS	BM808820		1071 bp mRNA linear EST 05-MAR-2002	
	AGENCOURT_5582944 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5471398		5', mRNA sequence.	
	BM808820		BM808820.1 GI:19125643	
	EST.		Homo sapiens (human)	
ACCESSION VERSION KEYWORDS SOURCE ORGANISM	Homo sapiens		Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	NIH-MGC http://mgi.nci.nih.gov/.		National Institutes of Health, Mammalian Gene Collection (MGC)	
	Unpublished		Contact: Robert Strausberg, Ph.D.	
REFERENCE AUTHORS TITLE JOURNAL COMMENT	Email: cgraps@email.nih.gov		Tissue Procurement: DCRD/DTF	
	cDNA Library Preparation: Rubin Laboratory		DNA Sequencing by: Agencourt Bioscience Corporation	
	found through the I.M.A.G.E. Consortium/LLNL at:		http://image.llnl.gov	
	Plate: L1CMI979 row: 1 column: 23		High quality sequence stop: 663.	
FEATURES source	1. .1071		/organism="Homo sapiens"	
	/mol_type="mRNA"		/db_xref="taxon:9606"	
	/clone="IMAGE:5471398"		/tissue_type="amelanotic melanoma, cell line"	
	/lab_host="DH10B (phage-resistant)"		/clone_lib="NIH_MGC_41"	

.....

DB 481 AACATGGAGTGTGGCACCCAGTCTTGGCGACAGTGCACAAAGTTTGGAGAGAACACC 540
QY 1516 TACTGCTTGGCGCGGACCGTGGAGACACTGCTGCTCGGCTTGGCAAGACCATCATGGAG 1575
DB 541 TACTGCTTGGCGCGGACCGTGGAGACACTGCTGCTCGGCTTGGCAAGACCATCATGGAG 600
QY 1576 GAGCAGCTGGTACTGAAGCGGGTGGCGCAACATCCTCATCAACCTGTATGGCATGACGGCC 1635
DB 601 GAGCAGCTGGTACTGAAGCGGGTGGCGCAACATCCTCATCAACCTGTATGGCATGACGGCC 660
QY 1636 GTGCTGTGCGGGCGGACCGCTCCATCCGCTATGGGCTCGGCAACACGACGACGAGTT 1695
DB 661 GTGCTGTGCGGGCGGACCGCTCCATCCGCTATGGGCTCGGCAACACGACGACGAGTT 720
QY 1696 CTCTTGCCCAACACCTTCTCGGGAAGCTTACTTGCAGAACTCTCTCAGCCTCTCTCAG 1755
DB 721 CTCTTGCGGCGACACCTTCTCGGGAAGCTTACTTGCAGAACTCTCTCAGCCTCTCTCAG 780
QY 1756 CTGCAAGATGCTCCAGAAACCTAGATGAGCAGATTAAAGAAAGTGTCCCGACGAGAT- 1814
DB 781 CTGCAAGATGCTCCAGAAACCTAGATGAGCAGATTAAAGAAAGTGTCCCGACGAGATC 840
QY 1815 CTTTGGAAGCGAGCGCTTATCTGTG-CCACACCTCTGGACAGGACATGC 1863
DB 841 CTTTGGAAGCGAGCGCTTATCTGTGCCCCACCTCTGGACAGGACATGC 890

RESULT 9
BM561207
LOCUS
DEFINITION
AGENCOURT_6589858 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5475994
5', mRNA sequence.
BM561207
BM561207.1 GI:18806273
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1079)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov.
Tissue Procurement: DCTD/DP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM1991 row: 1 column: 11
High quality sequence stop: 700.
Location/Qualifiers
1..1079
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5475994"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NTH_MGC_41"
/note="Organ: skin; Vector: pORB7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

295 a 245 c 311 g 226 t 2 others

BASE COUNT
ORIGIN

Query Match 45.4%; Score 845.4; DB 12; Length 1079;
Best Local Similarity 96.0%; Pred. No. 8.3e-221;
Matches 910; Conservative 0; Mismatches 32; Indels 6; Gaps 4;
QY 1 ATGAGCGGCTGGGGGCTCTTCTTGGCGACAGGCTGCGGCTGCTGCGCGGGGTCFG 60
DB 15 ATGAGCGGCTGGGGGCTCTTCTTGGCGACAGGCTGCGGCTGCTGCGCGGGGTCFG 74
QY 61 GTGGTCTCTACCGCGCAACCGCGGCTTACTTGCACACAGCCGCTGTAGAGCTTTCCGC 120
DB 75 GTGGTCTCTACCGCGCAACCGCGGCTTACTTGCACACAGCCGCTGTAGAGCTTTCCGC 134
QY 121 AAAGAGCTTTTCTTGGGCAAAATCAAGAAAGAAAGTGTTCCTTCCCAATTTCCAGAACTTAGC 180
DB 135 AAAGAGCTTTTCTTGGGCAAAATCAAGAAAGAAAGTGTTCCTTCCCAATTTCCAGAACTTAGC 194
QY 181 CAAGATGAACCTTAATGAATCAATCAGTCTTGGGACCGCTGGGAAATTTTGGAGAAA 240
DB 195 CAAGATGAACCTTAATGAATCAATCAGTCTTGGGACCGCTGGGAAATTTTGGAGAAA 254
QY 241 GAGGTGGACTCCCGAAATTTGACAGGAAGGAAATCCCAGATGAAATTTTGGAGAAA 300
DB 255 GAGGTGGACTCCCGAAATTTGACAGGAAGGAAATCCCAGATGAAATTTTGGAGAAA 314
QY 301 TTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCAGAAAGATATGTGGGCTGGGCTTC 360
DB 315 TTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCAGAAAGATATGTGGGCTGGGCTTC 374
QY 361 TCCACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGCTCATCATGTG 420
DB 375 TCCACACCATGTACTCAGACTAGGGGAGATCATCAGCATGGATGGGCTCATCATGTG 434
QY 421 ACCCTGGCAGCGCACCGCTATTGGGCTCAAGGGGATCATCTTGGCTGGGCTGAGGAG 480
DB 435 ACCCTGGCAGCGCACCGCTATTGGGCTCAAGGGGATCATCTTGGCTGGGCTGAGGAG 494
QY 481 CAGAAAGCAATTAATTCCTTAACTGCGCTCCCGGGAGACATTCAGAGCTTCTGGCTC 540
DB 495 CAGAAAGCAATTAATTCCTTAACTGCGCTCCCGGGAGACATTCAGAGCTTCTGGCTC 554
QY 541 ACGGAGCGACCGCTAGTGGGAGGATGCGAGCTCAATCCGGAGCAGACACACTAAGTGA 600
DB 555 ACGGAGCGACCGCTAGTGGGAGGATGCGAGCTCAATCCGGAGCAGACACACTAAGTGA 614
QY 601 GACAAAGCACTACATCTCAATGGCTTCAAGGCTCTGGATTAATGAGGACTGGCC 660
DB 615 GACAAAGCACTACATCTCAATGGCTTCAAGGCTCTGGATTAATGAGGACTGGCC 674
QY 661 AATATTTTACTGTGTTTGAAGACATGAGGCTGCTTCTGATTCGATGATCAGTGAAGAC 720
DB 675 AATATTTTACTGTGTTTGAAGACATGAGGCTGCTTCTGATTCGATGATCAGTGAAGAC 734
QY 721 AAAATCAGCATTCATAGTAGAAAGAGCTTTGGTGGAGTCACTAATGGAAACCCGAA 780
DB 735 AAAATCAGCATTCATAGTAGAAAGAGCTTTGGTGGAGTCACTAATGGAAACCCGAA 794
QY 781 GATAAATTAGCATTC-GGGGCTCCACACACTTGTGA- GTCCATTTTGAACACCAAGA 838
DB 795 GATAAATTAGCATTCGGGGCTCCACACACTTGTGAAGGTCCTATTTTGAACACCAAGA 854
QY 839 TACTGTGGAAACATCTTGTG-AGAGGTGGAGATGGGTAA- -GGTGGCCATGAC 894
DB 855 TACTGTGGAAACATCTTGGNAAAGGTCGGANATGGGTAAAGTGGGCCATGGAAC 914
QY 895 ATCTTCAACAGCGCGGTTTACAGCTGGCAGCGCTCTGGCTGGGCTG 942
DB 915 ATCTTCAACAGCGCGGTTTAAAGCATGGGCGAGCGGCTCTGGGCTG 962

RESULT 10
AL549834
LOCUS

AL549834 969 bp mRNA linear EST 31-MAY-2003

DEFINITION AL549834 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI054Y017 5-PRIME, mRNA sequence.

ACCESSION AL549834

VERSION AL549834.2

KEYWORDS GI:31271652

SOURCE EST.

ORGANISM Homo sapiens (human)

REFERENCE Homo sapiens

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE 1 (bases 1 to 969)

JOURNAL Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

COMMENT Full-length cDNA libraries and normalization

On Feb 15, 2001 this sequence version replaced gi:12886203.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 2225.f For

more information about this cluster, see

http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DI054AH09QPL&cluster=2225.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Paradise Avenue Genoscope sequence ID : CS0DI054AH09QPL.

Location/Qualifiers

1. 969

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DI054Y017"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 274 a 215 c 203 t 8 others

ORIGIN

Query Match 44.08; Score 819.6; DB 9; Length 969;

Best Local Similarity 98.78; Pred. No. 9.6e-214; Gaps 2;

Matches 841; Conservative 5; Mismatches 2; Indels 2;

1 ATGAGCGGCTGCGG-GCTCTTCCTGCGCACCACCGCTGCGGCTGCTGCGGGGTCT 59

119 ATGAGCGGCTGCGGNGCTCTCTGCGCACCACCGCTGCGGCTGCTGCGGGGTCT 178

60 GGTGGTCTCTACCGGACCGGGCTTACTCGCCAGCAGCCCGCTGAGAGCTTCGC 119

179 GGTGGTCTCTACGCGAACCGGAGGATCTCGCCAGCAGCCCGCTGAGAGCTTCGC 238

120 CAAGAGGCTTTCTAGGCAAAATCAAGAGAAAGAGTTTTCCTTCCATTTCCAGAGTAG 179

239 CAAGAGGCTTTCTAGGCAAAATCAAGAGAAAGAGTTTTCCTTCCATTTCCAGAGTAG 298

180 CCAAGATGAACCTTAATGAATCAATCACTGTTCTGGACCCCGTGGAAAATCTTCACGTG 239

299 CCAAGATGAACCTTAATGAATCAATCACTGTTCTGGACCCCGTGGAAAATCTTCACGTG 358

240 AGAGTGACCTCCGCAAAATGACCGAGGAGGAAATCCAGATGAACCTTTGGAGAA 299

359 AGAGTGACCTCCGCAAAATGACCGAGGAGGAAATCCAGATGAACCTTTGGAGAA 418

300 ATTGAAGAGCCTAGGCGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCTTGGGCTT 359

419 ATTGAAGAGCCTAGGCGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCTTGGGCTT 478

360 CTCACACCATGCTACTCAAGACTAGGGAGATCATCAGATGATGGTGGTCCATCACTGT 419

479 CTCACACCATGCTACTCAAGACTAGGGAGATCATCAGATGATGGTGGTCCATCACTGT 538

QY 420 GACCCTGGCAGCGCACCAGGCTATTGGCCTCAAGGGGATCATCTTGGCTGGCACTGAGGA 479

Db 539 GACCCTGGCAGCGCACCAGGCTATTGGCCTCAAGGGGATCATCTTGGCTGGCACTGAGGA 598

QY 480 GCAGAAAGCCAAATACCTTGCCTAAACTGGCTCCGGGAGCAGACATTCAGGCTTCTGCT 539

Db 599 GCAGAAAGCCAAATACCTTGCCTAAACTGGCTCCGGGAGCAGACATTCAGGCTTCTGCT 658

QY 540 CACGAGGCGCAGCAGCTGGGAGGATGCAGCCCTCAATCGGAGCAGACGACACTAAAGTGA 599

Db 659 CAAGGAGCCAGCAGCTGGGAGGATGCAGCCCTCAATCGGAGCAGACGACACTAAAGTGA 718

QY 600 AGACAAGAGCACTACATCTCCTCAATGGCTCCAAAGTCTGGATTACTATGAGGAGTGGC 659

Db 719 AGACAAGAGCACTACATCTCCTCAATGGCTCCAAAGTCTGGATTACTATGAGGAGTGGC 778

QY 660 CAATATTTTACTGTTTGCAGAACTGAGCTCGTGTGATTTCTGATGGATCAGTGAAGA 719

Db 779 CAATATTTTACTGTTTGCAGAACTGAGCTCGTGTGATTTCTGATGGATCAGTGAAGA 838

QY 720 CAAATATCAGCAGATTCATAGTAGAAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGA 779

Db 839 CAAATATCAGCAGATTCATAGTAGAAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGA 898

QY 780 AGATAAATTAGCATTCGGGGCTCCACACTTTGTGAAGTCCATTTTGAAGAACCAAGAT 839

Db 899 AGATAAATTAGCATTCGGGGCTCCACACTTTGTGAAGTCCATTTTGAAGAACCAAGAT 957

QY 840 ACCTGTGGAAAA 851

Db 958 ACCTGTGGAAAA 969

RESULT 11

CA488487

LOCUS

DEFINITION

AGENCOURT_10808657 MAPcL Homo sapiens cDNA clone IMAGE:6720247 5', mRNA sequence.

ACCESSION CA488487

VERSION CA488487.1

KEYWORDS GI:24950504

SOURCE EST.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 884)

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Kristi A. Eglund, Ira Pastan

cDNA Library Preparation: Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAMA4279 row: h column: 07

High quality sequence stop: 713.

Location/Qualifiers

1. 884

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6720247"

/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1, LNCaP"

/lab_host="EMDH10B"

/clone_lib="MAPCL"

/note="vector: pCMV-SPORT6; Site_1: EcorV; Site_2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dT. Average

FEATURES

source

Insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglund, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted.

```
BASE COUNT      219 a   219 c   260 g   186 t
ORIGIN
Query Match      43.5%; Score 810.6; DB 14; Length 884;
Best Local Similarity 98.8%; Pred. No. 2.7e-211;
Matches 827; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
QY 695 TTGATCTCATGATGAGTGAAGACAAAATCACAGCATTCATAGTAGAAGAGACTTTG 754
Db 1 TTGATCTCATGATGAGTGAAGACAAAATCACAGCATTCATAGTAGAAGAGACTTTG 60
QY 755 GTGGAGTCACTAATGGAAACCCGAGATTAATAGGCATTCGGGCTCCACACTTTG 814
Db 61 GTGGAGTCACTAATGGAAACCCGAGATTAATAGGCATTCGGGCTCCACACTTTG 120
QY 815 AAGTCCATTTTGAACACCAAGATACCTGTGAAACATCCTTTGGAGAGTGGAGATG 874
Db 121 AAGTCCATTTTGAACACCAAGATACCTGTGAAACATCCTTTGGAGAGTGGAGATG 180
QY 875 GGTTTAAGTGGCCATGAACATCCTCAACAGCGCGGTTTCAGCATGGCAGCGCTGG 934
Db 181 GGTTTAAGTGGCCATGAACATCCTCAACAGCGCGGTTTCAGCATGGCAGCGCTGG 240
QY 935 CTGGGCTGCTCAAGAGATGATTGAATGACCTGCTGAGTACCCCTGCACAGAAACAGT 994
Db 241 CTGGGCTGCTCAAGAGATGATTGAATGACCTGCTGAGTACCCCTGCACAGAAACAGT 300
QY 995 TTAAACAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCACCTGATGGCTCAGA 1054
Db 301 TTAAACAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCACCTGATGGCTCAGA 360
QY 1055 AGGCTTACGTCATGAGAGATGATGACCTACCTCAGACGAGGATGCTGGACCAACCTGGCT 1114
Db 361 AGGCTTACGTCATGAGAGATGATGACCTACCTCAGACGAGGATGCTGGACCAACCTGGCT 420
QY 1115 TTCCGAGCTGCTCCATCGAGGAGGAGCATGTTGAAGTGTTCAGCTCCGAGCGCGCTGGC 1174
Db 421 TTCCGAGCTGCTCCATCGAGGAGGAGCATGTTGAAGTGTTCAGCTCCGAGCGCGCTGGC 480
QY 1175 AGTGTGTGAGTGAGGCGCTGCAGATCTCGGGGGCTTGGGCTTACAAAGGAGTATCCGT 1234
Db 481 AGTGTGTGAGTGAGGCGCTGCAGATCTCGGGGGCTTGGGCTTACAAAGGAGTATCCGT 540
QY 1235 ACGAGCGCATACTGCTGACACCCGCTCCTCCTCATCTTCAGGAGAACCAATGAGATTC 1294
Db 541 ACGAGCGCATACTGCTGACACCCGCTCCTCCTCATCTTCAGGAGAACCAATGAGATTC 600
QY 1295 TCCGATGATACATCCCTGACGGGTCTGCAGCATGCCGGCGCATCTCCTGACTACACAGA 1354
Db 601 TCCGATGATACATCCCTGACGGGTCTGCAGCATGCCGGCGCATCTCCTGACTACACAGA 660
QY 1355 TCCATGACCTTAAACAGGCCAAAGTAGACACAGTCAATGATACCGTTGGCCGGAGGCTTC 1414
Db 661 TCCATGACCTTAAACAGGCCAAAGTAGACACAGTCAATGATACCGTTGGCCGGAGGCTTC 720
QY 1415 GGGACTCCTCGGGCGGAACCTGTGGACCTGGGGCTGCACAGCAACCATGGAGTTGGACACC 1474
Db 721 GGGACTCCTCGGGCGGAACCTGTGGACCTGGGGCTGCACAGCAACCATGGAGTTGGACACC 780
QY 1475 CCAGTCTTGGCGACAGTGCCCAACAG-TTTGAGGAGAACACCTACTGCTTCGGCGG 1530
Db 781 CCAGTCTTGGCGACAGTGCCCAACAGTTTTCAGGGAGACACCTACTGCTTCGGCGG 837
RESULT 12
BU528453
LOCUS
DEFINITION AGENCOURT_1018068 NIH_MGC_101 Homo sapiens cDNA clone
EST 13-SEP-2002
```

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

IMAGE:6538399 5', mRNA sequence.
BU528453
BU528453.1 GI:22838894
EST.

Homo sapiens (human)
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia: Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 922)
NIH-MGC http://imgc.ncbi.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LLM2702 row: o column: 07
High quality sequence start: 6
High quality sequence stop: 613.
Location/Qualifiers
1. 922

FEATURES
source

1. 922

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6538399"

/tissue_type="epidermoid carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_101"

/note="Organ: lung; Vector: pORF7; Site:1: EcoRI; Site:2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 208 a 256 c 270 g 187 t 1 others

ORIGIN

Query Match 43.5%; Score 810; DB 13; Length 922;
Best Local Similarity 96.5%; Pred No. 4.1e-211;
Matches 881; Conservative 0; Mismatches 26; Indels 6; Gaps 5;

QY 798 GGGCTCCACACTTGTGAAGTCCATTTTGAACACCAAGATACCTGTGGAACACATCCT 857

Db 1 GGGCTCCACACTTGTGAAGTCCATTTTGAACACCAAGATACCTGTGGAACAC -TCCT 59

QY 858 TGGAGAGTCCGAGATGGGTTTAAGTGGCCATGAACATCTCAACAGCGCGGCTTCAG 917

Db 60 TGGAGAGTCCGAGATGGGTTTAAGTGGCCATGAACATCTCAACAGCGCGGCTTCAG 119

QY 918 CATGGGAGCGCTGCTGGCTGGCTGCTCAAGAGATTGATTGAATGACTGCTGAGTAGGC 977

Db 120 CATGGGAGCGCTGCTGGCTGGCTGCTCAAGAGATTGATTGAATGACTGCTGAGTAGGC 179

QY 978 CTGCACAGGAAACAGTTTAAACAGAGCTCAGTCAATTTGGATTGATTCAGGAGAAAT 1037

Db 180 CTGCACAGGAAACAGTTTAAACAGAGCTCAGTCAATTTGGATTGATTCAGGAGAAAT 239

QY 1038 TGCACTGATGCTCAGAGGCTTACGTCATGGAGAGTATGACCTACCTCAGCAGGAGAT 1097

Db 240 TGCACTGATGCTCAGAGGCTTACGTCATGGAGAGTATGACCTACCTCAGCAGGAGAT 299

QY 1098 GCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGCAGCCATGTTGAAGTGTTCAG 1157

Db 300 GCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGCAGCCATGTTGAAGTGTTCAG 359

QY 1158 CTCCGAGCGCGCTGGCAGTGTGTGAGTGAGGCGCTGCAGATCTCGGGGCTTGGGCTA 1217

```
|||||
360 CTCGAGCGCGCTGCGAGTGTGAGTGAGGCGCTGCAGATCCTCGGGGCTTGGGCTA 419
QY 1218 CACAAGGAGTATCCGTAGAGGGGATGCTGGTGACACCGGATCCTCTCATCTTCA 1277
Db 420 CACAAGGAGTATCCGTAGAGGGGATGCTGGTGACACCGGATCCTCTCATCTTCA 479
QY 1278 GGAACCAATGAGATTCGCGATGTACATCGCCTGACGGGTCTGCAGATCGCGGCG 1337
Db 480 GGAACCAATGAGATTCGCGATGTACATCGCCTGACGGGTCTGCAGATCGCGGCG 539
QY 1338 CATCTGTACTACAGGATCCATGAGCTTAAACAGGCGCAAGTGGAGCAGTATGATAC 1397
Db 540 CATCTGTACTACAGGATCCATGAGCTTAAACAGGCGCAAGTGGAGCAGTATGATAC 599
QY 1398 CGTTGGCGGAGGCTTCGGAGTCCCTGGGCGGAAGTGGAGCAGTATGATAC 1457
Db 600 CGTTGGCGGAGGCTTCGGAGTCCCTGGGCGGAAGTGGAGCAGTATGATAC 659
QY 1458 CCATGAGTGTGACACCGGAGTCTGCGGACAGTGCACAAAGTTTGGAGGAGAACCTA 1517
Db 660 CCATGAGTGTGACACCGGAGTCTGCGGACAGTGCACAAAGTTTGGAGGAGAACCTA 719
QY 1518 CTGCTTGGCGGAGGCTGGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1577
Db 720 CTGCTTGGCGGAGGCTGGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
QY 1578 GCAGCTGTGACTGAAGC-GGGTGGCAACATCTCATCAACCTGATGATGATGATGAT 1635
Db 780 GCAGCTGTGACTGAAGC-GGGTGGCAACATCTCATCAACCTGATGATGATGATGAT 839
QY 1636 GTGCTGTGCGGGGCGGAG-CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1692
Db 840 GTGCTGTGCGGGGCGGAG-CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899
QY 1693 GTTCTCTTGGGCA 1705
Db 900 GTTCTCTTGGGCA 912

RESULT 13
LOCUS BUI76038
DEFINITION AGENCOURT_7782617 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6137085
5', mRNA sequence.
ACCESSION BUI76038
VERSION BUI76038.1 GI:22690022
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-i@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13450 row: e column: 22
High quality sequence stop: 651.
Location/Qualifiers
1..851
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6137085"

FEATURES
source
```

```
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT 234 a 205 c 230 g 181 t 1 others
ORIGIN
Query Match 43.0%; Score 800.4; DB 13; Length 851;
Best Local Similarity 99.4%; Pred. No. 1.7e-208;
Matches 834; Conservative 0; Mismatches 2; Indels 3; Gaps 3;
QY 1 ATGAGCGGCTGGGGGCTCTTCTGCGCACACGCGCTGCGCTGCGCTGCGCTGCGCTGCG 60
Db 14 ATGAGCGGCTGGGGGCTCTTCTGCGCACACGCGCTGCGCTGCGCTGCGCTGCGCTGCG 73
QY 61 GTGGTCTTACCGCGAACCGCGGCTTACTGCGCACCGCGCTGCGCTGCGCTGCGCTTTCG 120
Db 74 GTGGTCTTACCGCGAACCGCGGCTTACTGCGCACCGCGCTGCGCTGCGCTTTCGCTC 133
QY 121 AAGAGCTTTTCTAGCGAAATCAAGAAAGAAAGTTCCTCCATTTCCAGAAAGTAGC 180
Db 134 AAGAGCTTTTCTAGCGAAATCAAGAAAGAAAGTTCCTCCATTTCCAGAAAGTAGC 193
QY 181 CAAGATGAAGTAAATGAAATCAATCAAGTTCCTGGGACCGCTGGGAAATTTCTTCACTGA 240
Db 194 CAAGATGAAGTAAATGAAATCAATCAAGTTCCTGGGACCGCTGGGAAATTTCTTCACTGA 253
QY 241 GAGGTGGACTCCGGAATAATGACAGGAGGAGAAATCCAGATCAAACTTTGGAGAA 300
Db 254 GAGGTGGACTCCGGAATAATGACAGGAGGAGAAATCCAGATCAAACTTTGGAGAA 313
QY 301 TTGAAGAGCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCTTGGGCTTC 360
Db 314 TTGAAGAGCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCTTGGGCTTC 373
QY 361 TCCACACCATGCTACTCAAGACTTAGGGGAGATCATCAGCATGGATGGGCTCATCTGTG 420
Db 374 TCCACACCATGCTACTCAAGACTTAGGGGAGATCATCAGCATGGATGGGCTCATCTGTG 433
QY 421 ACCCTGGCAGCGCAGCAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCAGTGGAG 480
Db 434 ACCCTGGCAGCGCAGCAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCAGTGGAG 493
QY 481 CAGAAAGCCAAATACTTGGCTAAACTGGGCTCCGGGAGCAGCATTCAGGCTTCTGCCTC 540
Db 494 CAG-AAGCCAAATACTTGGCTAAACTGGGCTCCGGGAGCAGCATTCAGGCTTCTGCCTC 552
QY 541 ACGGAGCCAGCGAGTGGGAGCGATGAGCGCTCAATCCGGAGCAGAGCCACACTAAGTGA 600
Db 553 ACGGAGCCAGCGAGTGGGAGCGATGAGCGCTCAATCCGGAGCAGAGCCACACTAAGTGA 612
QY 601 GACAAGCAGCTACATCTCTCAATGCTCCAGGCTCTGGATTACTAATGAGGAGCTGGCC 660
Db 613 GACAAGCAGCTACATCTCTCAATGCTCCAGGCTCTGGATTACTAATGAGGAGCTGGCC 672
QY 661 AATATTTTACTGTGTGTCAAAGACTGAGGCTGCTGATTTCTGATGGATGAGTGAAGAC 720
Db 673 AATATTTTACTGTGTGTCAAAGACTGAGGCTGCTGATTTCTGATGGATGAGTGAAGAC 732
QY 721 AATATCAGCATTTTACTAGTAAAGAGACTTTGGTGGAGTCACTAAT-GGGAAACCCGA 779
Db 733 AATATCAGCATTTTACTAGTAAAGAGACTTTGGTGGAGTCACTAATGGGGAACCCGA 792
QY 780 AGATAAATAGGATTC-GGGGCTCCAAACACTTGTGAAGTCCATTTTGAAGACCAAG 837
Db 793 AGATAAATAGGATTCGGGGGCTCCAAACACTTGTGAAGTCCATTTTGAAGACCAAG 851

RESULT 14
BQ960770
```

```

LOCUS      BQ960770      888 bp      mRNA      linear      EST 21-AUG-2002
DEFINITION AGENCOURT_8881366 Lupski_sciatic_nerve Homo sapiens cDNA clone
IMAGE:6201350 5', mRNA sequence.
BQ960770
VERSION    BQ960770.1 GI:22376248
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 888)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13617 row: k column: 15
High quality sequence stop: 640.
Location/Qualifiers
1..888
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6201350"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="Lupski_sciatic_nerve"
Note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dr priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGGCTCCG-3' and
5'-GACTAGTTCATGATCGGAGCGGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
BASE COUNT 252 a 194 c 250 g 192 t
ORIGIN

Query Match      42.9%; Score 798.8; DB 13; Length 888;
Best Local Similarity 97.8%; Pred. No. 4.8e-208;
Matches 833; Conservative 0; Mismatches 12; Indels 7; Gaps 2;

QY 194 ATGAATCAATCAGTCTTGGGACCCCTGGAAAAATTTCTTCACTGAAGAGGTGGACATCCC 253
Db 12 ATGAATCAATCAGTCTTGGGACCCCTGGAAAAATTTCTTCACTGAAGAGGTGGACATCCC 71
QY 254 GAAAAATTGACACGAGGGAATCCAGATGAACCTTTGGAGAAATTTGAAGAGCCCTAG 313
Db 72 GAAAAATTGACACGAGGGAATCCAGATGAACCTTTGGAGAAATTTGAAGAGCCCTAG 131
QY 314 GCGTTTTGGGCTCCAAAGTCCCAAGAAATATGTTGGCCCTGGGCTTCTCCAAACACCATGT 373
Db 132 GCGTTTTGGGCTCCAAAGTCCCAAGAAATATGTTGGCCCTGGGCTTCTCCAAACACCATGT 191
QY 374 ACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCATCTGTACCCCTGGCAGCGC 433
Db 192 ACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCATCTGTACCCCTGGCAGCGC 251
QY 434 ACCAGGCTATTGGCCTCAAGGGGATCATCTTGGCTGCGACTGAGGAGCAAGGCCAAAT 493
Db 252 ACCAGGCTATTGGCCTCAAGGGGATCATCTTGGCTGCGACTGAGGAGCAAGGCCAAAT 311

```

```

QY 494 ACTTGCCTAACTGGCTCCGGGAGACACATTTGCAGCCTTCTGCCTCAGGAGCCAGCCA 553
Db 312 ACTTGCCTAACTGGCTCCGGGAGACACATTTGCAGCCTTCTGCCTCAGGAGCCAGCCA 371
QY 554 GTGGGAGCGATGCAGCCTCAATCCGGAGCAGAGCCACACTAAGTCAAGACAAAGACACT 613
Db 372 GTGGGAGCGATGCAGCCTCAATCCGGAGCAGAGCCACACTAAGTCAAGACAAAGACACT 431
QY 614 ACATCCTCAATGGCTCCCAAGGCTCTGGATTACTTAATGGAGGACTGGCCCAATATTTTACTG 673
Db 432 ACATCCTCAATGGCTCCCAAGGCTCTGGATTACTTAATGGAGGACTGGCCCAATATTTTACTG 491
QY 674 TGTTTGCAAGACTGAGGCTCGTTGATTTCTGATGATGATGATGATGATGATGATGATGAT 733
Db 492 TGTTTGCAAGACTGAGGCTCGTTGATTTCTGATGATGATGATGATGATGATGATGATGAT 551
QY 734 TCATAGTAGAAGAGACTTTTGTGTGAGTCACTAATGGGAAACCCGAGAGATAAATAGGCA 793
Db 552 TCATAGTAGAAGAGACTTTTGTGTGAGTCACTAATGGGAAACCCGAGAGATAAATAGGCA 611
QY 794 TTCGGGGCTCAACACACTTGTGAAGTCCATTTTGAAACACCAAGATACCTGTGGAAACA 853
Db 612 TTCGGGGCTCAACACACTTGTGAAGTCCATTTTGAAACACCAAGATACCTGTGGAAACA 671
QY 854 TCCTTTGGAGAGTCGGAGATGGGTTTAAGTGGCCATGAACATCCTCAACAGCGGCCGT 913
Db 672 TCCTTTGGAGAGTCGGAGATGGGTTTAAGTGGCCATGAACATCCTCAACAGCGGCCGT 731
QY 914 TCAGCATGGGAGCGCTCGTGGCTGGCTGCTCAAGAGATTTGATTTTGGTAGAATGACTGCT 969
Db 732 TCAGCATGGGAGCGCTCGTGGCTGGCTGCTCAAGAGATTTGATTTTGGTAGAATGACTGCT 791
QY 970 GAGTACGCTCCCAACAGGAAACAGTTTAAACAGAGGCTCAGTGAA---TTTGATTTGATT 1026
Db 792 GAGTACGCTCCCAACAGGAAACAGTTTAAACAGAGGCTCAGTGAA---TTTGATTTGATT 851
QY 1027 CAGGAGAAATTT 1038
Db 852 CCGGAGAAATTT 863

```

```

RESULT 15
BQ938039
LOCUS
DEFINITION AGENCOURT_8948875 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6469922
5', mRNA sequence.
BQ938039
VERSION    BQ938039.1 GI:22353517
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 890)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13999 row: n column: 03
High quality sequence stop: 611.
Location/Qualifiers
1..890
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```

FEATURES
source

Tue Sep 9 10:14:39 2003

us-09-945-326-3.rst

```

/clone="IMAGE:6469922"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."
BASE COUNT      242 a      214 c      248 g      186 t
ORIGIN
Query Match      42.9%; Score 798.4; DB 13; Length 890;
Best Local Similarity 99.5%; Pred. No. 6.2e-208;
Matches 832; Conservative 0; Mismatches 1; Indels 3; Gaps 3;
Qy 1 ATGAGCGGCTGCGGGGCTCTTCCTGCGCACCACCGCTGCGGCTCGTGCCTGCCGCGGGTCTG 60
Db 55 ATGAGCGGCTGCGGGGCTCTTCCTGCGCACCACCGCTGCGGCTCGTGCCTGCCGCGGGTCTG 114
Qy 61 GTGGTCTCTACCGCGAACCGCGGCTACTGCGCACCACCGCGGCTGTAGGAGCTTTTCGCC 120
Db 115 GTGGTCTCTACCGCGAACCGCGGCTACTGCGCACCACCGCGGCTGTAGGAGCTTTTCGCC 174
Qy 121 AAAGAGCTTTTCTTAGGCAAAATCAAGAGAAAGAGTTTTCCTCATTTCCAGAGTAGC 180
Db 175 AAAGAGCTTTTCTTAGGCAAAATCAAGAGAAAGAGTTTTCCTCATTTCCAGAGTAGC 234
Qy 181 CAAGATCAACTTAATGAATCAATCAGTTCTTGGGACCGCTGGAAAAATTCCTCACTGAA 240
Db 235 CAAGATCAACTTAATGAATCAATCAGTTCTTGGGACCGCTGGAAAAATTCCTCACTGAA 294
Qy 241 GAGGTGAGCTCCCGAAAAATGACAGGAGGAAAAATCCAGATGAATTTGGAGAAA 300
Db 295 GAGGTGAGCTCCCGAAAAATGACAGGAGGAAAAATCCAGATGAATTTGGAGAAA 354
Qy 301 TTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCCAGAGAAATATGTTGGCTGGGCTTC 360
Db 355 TTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCCAGAGAAATATGTTGGCTGGGCTTC 414
Qy 361 TCCAAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCACTGTG 420
Db 415 TCCAAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCACTGTG 474
Qy 421 ACCCTGGCAGCGCACCGAGCTATTGGGCTCAAGGGGATCATCTTGGCTGGCAGTGGAG 480
Db 475 ACCCTGGCAGCGCACCGAGCTATTGGGCTCAAGGGGATCATCTTGGCTGGCAGTGGAG 534
Qy 481 CAGAAAGCCAAATACTTGCCTFAAACTGGCGTCCGGGGAGACATTCAGAGCTTCTGCCTC 540
Db 535 CAGAAAGCCAAATACTTGCCTFAAACTGGCGTCCGGGGAGACATTCAGAGCTTCTGCCTC 594
Qy 541 ACGGAGCCAG-CCAGTGGGAGCGATGAGCGCTCAATCCGGAGCAGAGCCACACTAAGTGA 599
Db 595 ACGGAGCCAGCCAGTGGGAGCGATGAGCGCTCAATCCGGAGCAGAGCCACACTAAGTGA 654
Qy 600 AGACAAGAGCACTACATCTCAATGGCTCCCAAGTCTGGAATTAATGAGGAGCTGGC 659
Db 655 AGACAAGAGCACTACATCTCAATGGCTCCCAAGTCTGGAATTAATGAGGAGCTGGC 714
Qy 660 CAATATTTTACTGTGTTTGAAGAGCTGAGTCTGATTTCTGATGATCAGTGAAGA 719
Db 715 CAATATTTTACTGTGTTTGAAGAGCTGAGTCTGATTTCTGATGATCAGTGAAGA 774
Qy 720 CAAATACAGCAATCATAGTAGAAGAGACTTTTGGTGGAGTCAATGGAAGACCCG 778
Db 775 CAAATACAGCAATCATAGTAGAAGAGACTTTGGTGGAGTCAATGGAAGACCCG 834
Qy 779 AAGATAAATAGGCAATTCGGGCT-CCAAACACTTGTGAAGTCCATTTTGAACAC 833
Db 835 AAGATAAATAGGCAATTCGGGCTCCCAACACTTGTGAGTCCATTTTGAACAC 890

```

Search completed: September 7, 2003, 01:51:44
Job time : 3922.86 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2003, 23:17:14 ; Search time 421.388 Seconds
(without alignments)
10168.445 Million cell updates/sec

Title: US-09-945-326-3
Perfect score: 1863
Sequence: 1 atgagcggtcgggctctt.....accctctggacgacatgc 1863

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1537136 seqs, 114998732 residues
Total number of hits satisfying chosen parameters: 3074272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:**
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Description	
Result No.	Score	Query Match	Length	ID		
1	1863	100.0	1863	10	US-09-945-326-3	Sequence 3, Appl1
2	1863	100.0	2452	10	US-09-945-326-1	Sequence 1, Appl1
3	1861.4	99.9	2440	14	US-10-168-274-51	Sequence 51, Appl1
4	1064.6	57.1	2080	14	US-10-198-846-11756	Sequence 11756, A
5	452.2	24.3	490	11	US-09-918-995-21901	Sequence 21901, A
6	451.8	24.3	483	11	US-09-918-995-21907	Sequence 21907, A
7	390.6	21.0	451	11	US-09-918-995-23707	Sequence 16927, A
8	388.6	20.9	2117	10	US-09-917-800A-11034	Sequence 11034, A
9	261.4	14.0	490	11	US-09-918-995-16927	Sequence 1584, Ap
10	214.6	11.5	312	9	US-09-822-849A-480	Sequence 16927, A
11	165.2	8.9	568	10	US-09-756-692-8175	Sequence 480, App
12	165.2	8.9	568	14	US-10-040-862-8175	Sequence 8175, Ap
13	159.4	8.6	183	10	US-09-867-701-3392	Sequence 3392, Ap
14	151.2	8.1	1158	14	US-10-156-761-5259	Sequence 5259, Ap
15	151.2	8.1	9025608	14	US-10-156-761-1	Sequence 1, Appli
16	147.6	7.9	1170	14	US-10-156-761-6897	Sequence 6897, Ap

17	147.6	7.9	9025608	14	US-10-156-761-1	Sequence 1, Appli
18	147.4	7.9	302	11	US-09-918-995-20366	Sequence 20366, A
19	146.6	7.9	1829	10	US-09-880-107-2309	Sequence 2309, Ap
20	137.8	7.4	402	11	US-09-918-995-4541	Sequence 4541, Ap
21	132.6	7.1	1155	14	US-10-156-761-5010	Sequence 5010, Ap
22	126.6	6.8	1716	10	US-09-974-300-1101	Sequence 1101, Ap
23	120.8	6.5	1138	10	US-09-974-300-1082	Sequence 1082, Ap
24	119.8	6.4	1149	14	US-10-156-761-6577	Sequence 6577, Ap
25	118.2	6.3	1866	10	US-09-917-800A-1617	Sequence 1617, Ap
26	115.6	6.2	810	10	US-09-974-300-5530	Sequence 5530, Ap
27	114.4	6.1	1155	14	US-10-156-761-5264	Sequence 5264, Ap
28	111.2	6.0	17417	14	US-10-272-419-27	Sequence 5, Appli
29	111.2	6.0	17417	14	US-10-272-419-27	Sequence 27, Appli
30	110.6	5.9	1056	10	US-09-974-300-5546	Sequence 5546, Ap
31	105	5.6	1938	14	US-10-156-761-2554	Sequence 2554, Ap
32	104.2	5.6	411	10	US-09-960-352-1781	Sequence 1781, Ap
33	103.6	5.6	439	10	US-09-918-995-317	Sequence 317, App
34	102.4	5.5	495	11	US-09-918-995-350	Sequence 350, App
35	101.8	5.5	376	14	US-10-060-036-766	Sequence 766, App
36	100.2	5.4	562	10	US-09-998-598-1171	Sequence 1171, Ap
37	94.6	5.1	2238	14	US-10-205-823-3	Sequence 3, Appli
38	92.6	5.0	983	10	US-09-974-300-1090	Sequence 1090, Ap
39	90	4.8	2682	10	US-09-880-107-3300	Sequence 3300, Ap
40	90	4.8	2682	12	US-10-354-358-27	Sequence 27, Appl
41	81.8	4.4	615	10	US-09-974-300-1102	Sequence 1102, Ap
42	81.4	4.4	1383	10	US-09-925-300-705	Sequence 705, App
43	80.6	4.3	1203	14	US-10-156-761-1904	Sequence 1904, Ap
44	80	4.3	1798	10	US-09-880-107-3413	Sequence 3413, Ap
45	78.2	4.2	1125	14	US-10-156-761-2293	Sequence 2293, Ap

ALIGNMENTS

RESULT 1
US-09-945-326-3
; Sequence 3, Application US/09945326
; Patent No. US20020127680A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Hunter, John Joseph
; TITLE OF INVENTION: 62112, A NOVEL HUMAN DEHYDROGENASE AND
; FILE OF INVENTION: US THEREOF
; FILE REFERENCE: MNT-187
; CURRENT APPLICATION NUMBER: US/09/945,326
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/229,831
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1863
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(1863)
US-09-945-326-3

Query Match 100.0%; Score 1863; DB 10; Length 1863;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCGGCTCGGGCTCTTCTCGGCACACGCTCGGGCTCGTGGCTCGGCGGGTCTG 60
1 ATGAGCGGCTCGGGCTCTTCTCGGCACACGCTCGGGCTCGGCGGGTCTG 60
DB 1 ATGAGCGGCTCGGGCTCTTCTCGGCACACGCTCGGGCTCGTGGCTCGGCGGGTCTG 60
QY 61 GTGTGCTCTTACCGCGAACCGCGGCTACTGCGCACCGCCCTGTACGAGCTTTCGCC 120
1 GTGTGCTCTTACCGCGAACCGCGGCTACTGCGCACCGCCCTGTACGAGCTTTCGCC 120
DB 61 GTGTGCTCTTACCGCGAACCGCGGCTACTGCGCACCGCCCTGTACGAGCTTTCGCC 120
QY 121 AAAGAGCTTTTCCTAGGCAAAATCAAGAGAAAGAGTTTTCCTATTTCCAGAGTTAGC 180
1 AAAGAGCTTTTCCTAGGCAAAATCAAGAGAAAGAGTTTTCCTATTTCCAGAGTTAGC 180

Matches 1863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAGCGGCTGCGGGCTCTTCTCGCGCACACAGCGCTGCGGCTCGTGCCTTCGGGGTCTG	60
Db	67	ATGAGCGGCTGCGGGCTCTTCTCGCGCACACAGCGCTGCGGCTCGTGCCTTCGGGGTCTG	126
QY	61	GTGCTCTTACCAGAACCGGGCTTACTGCGCACAGCCGCCCTGTACGAGCTTTTCGCC	120
Db	127	GTGCTCTTACCAGAACCGGGCTTACTGCGCACAGCCGCCCTGTACGAGCTTTTCGCC	186
QY	121	AAAGAGCTTTTCTTAGSCAAATCAAGAAAGAAAGTTTTCCTTCCATTTCCAGAAGTTAGC	180
Db	187	AAAGAGCTTTTCTTAGSCAAATCAAGAAAGAAAGTTTTCCTTCCATTTCCAGAAGTTAGC	246
QY	181	CAAGATGAACCTTAATGAATCAATCAGTTCTTGGGACCGTGGAAAAATCTTCACTGAA	240
Db	247	CAAGATGAACCTTAATGAATCAATCAGTTCTTGGGACCGTGGAAAAATCTTCACTGAA	306
QY	241	GAGGTGACCTCCGAAAAATTGACAGGAGGAAAAATCCCAAGATGAAACTTTTGAGAAA	300
Db	307	GAGGTGACCTCCGAAAAATTGACAGGAGGAAAAATCCCAAGATGAAACTTTTGAGAAA	366
QY	301	TTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCACAGAGAAATAGTGGCTGGGCTTC	360
Db	367	TTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCACAGAGAAATAGTGGGCTGGGCTTC	426
QY	361	TCCAACACCATGACTCAAGACTTAGGGAGATCATCAGCATGGATGGTCCATCACATGTG	420
Db	427	TCCAACACCATGACTCAAGACTTAGGGAGATCATCAGCATGGATGGTCCATCACATGTG	486
QY	421	ACCTGGCAGCGACACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCACTGAGGAG	480
Db	487	ACCTGGCAGCGACACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCACTGAGGAG	546
QY	481	CAGAAAGCCAAATACTTGCTTAACTTGGCTCGGGGAGACATGCGAGCCTTCGGCTC	540
Db	547	CAGAAAGCCAAATACTTGCTTAACTTGGCTCGGGGAGACATGCGAGCCTTCGGCTC	606
QY	541	ACGGAGCCAGCCAGTGGGAGCGATGACGCTCAATCCGGAGCAGAGCCACACTAAAGTAA	600
Db	607	ACGGAGCCAGCCAGTGGGAGCGATGACGCTCAATCCGGAGCAGAGCCACACTAAAGTAA	666
QY	601	GACAAGACACTACATCCTCAATGGCTCCAAAGTCTGGATTACTAATGGAGGACTGGCC	660
Db	667	GACAAGACACTACATCCTCAATGGCTCCAAAGTCTGGATTACTAATGGAGGACTGGCC	726
QY	661	AATATTTTTACTGTGTTTGCAAGACTGAGGTCTGTGATTCTGATGATCAGTGAAGAC	720
Db	727	AATATTTTTACTGTGTTTGCAAGACTGAGGTCTGTGATTCTGATGATCAGTGAAGAC	786
QY	721	AAATCAGACGATTCATAGTAGAAAGACACTTTGGTGGAGTCACTAATGGGAAACCCGAA	780
Db	787	AAATCAGACGATTCATAGTAGAAAGACACTTTGGTGGAGTCACTAATGGGAAACCCGAA	846
QY	781	GATAAATTAGGCATTCGGGGCTCCAACACTTTGTGAAGTCCATTTTGAACACCAAGATA	840
Db	847	GATAAATTAGGCATTCGGGGCTCCAACACTTTGTGAAGTCCATTTTGAACACCAAGATA	

RESULT 3
US-10-168-274-51
: Sequence 51, Application US/10168274
: Publication No. US20030124106A1
: GENERAL INFORMATION:
: APPLICANT: INCYTE GENOMICS, INC.
: APPLICANT: YUE, Henry
: APPLICANT: LAL, Preeti
: APPLICANT: TANG, Y. Tom
: APPLICANT: HILLMAN, Jennifer
: APPLICANT: BAUGHN, Mariah R.
: APPLICANT: AZIMZAI, Yalda
: APPLICANT: LO, Dyoung Alina M.
: TITLE OF INVENTION: HUMAN OXIDOREDUCTASE PROTEINS
: FILE REFERENCE: PF-0754 PCT

; CURRENT APPLICATION NUMBER: US/10/168,274
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/172,367
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 51
; LENGTH: 2440
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030124106A1 5540437CB1
US-10-168-274-51

Query Match 99.9%; Score 1861.4; DB 14; Length 2440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1862; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGCGGCTGCGGGCTCTTCCCTGCGCACACAGGCTGCGCTCGTGCCTGCGGGGTCTG 60
DB 60 ATGAGCGGCTGCGGGCTCTTCCCTGCGCACACAGGCTGCGCTCGTGCCTGCGGGGTCTG 119
QY 61 GTGCTCTTACCCGCAACCGGGCTACTGCGCACACAGGCTGCGCTGTAGAGCTTTCGCC 120
DB 120 GTGCTCTTACCCGCAACCGGGCTACTGCGCACACAGGCTGCGCTGTAGAGCTTTCGCC 179
QY 121 AAGAGCTTTCTAGGCAAAATCAAGAGAAAGAAATTTCCCATTTCCAGAAATGAGC 180
DB 180 AAGAGCTTTCTAGGCAAAATCAAGAGAAAGAAATTTCCCATTTCCAGAAATGAGC 239
QY 181 CAAGATGAACCTTAATGAATCAATCAGTTCTTGGGACCGCTGGAAATTTCTCACTGAA 240
DB 240 CAAGATGAACCTTAATGAATCAATCAGTTCTTGGGACCGCTGGAAATTTCTCACTGAA 299
QY 241 GAGGTGACCTCCGAAATTTACAGAGAGAGAAATCCCAGATGAACCTTTGGAGAA 300
DB 300 GAGGTGACCTCCGAAATTTACAGAGAGAGAAATCCCAGATGAACCTTTGGAGAA 359
QY 301 TTGAAGAGCTAGGGCTTTTGGGCTCAAGTCCAGAGAAATATGTTGGCTGCGCTTC 360
DB 360 TTGAAGAGCTAGGGCTTTTGGGCTCAAGTCCAGAGAAATATGTTGGCTGCGCTTC 419
QY 361 TCCAAACCACTGACTCAAGCTAGGGAGATCATCAGCATGAGTGGTCCATCACTGTG 420
DB 420 TCCAAACCACTGACTCAAGCTAGGGAGATCATCAGCATGAGTGGTCCATCACTGTG 479
QY 421 ACCCTGCGAGCGCACAGGCTATTGGCTCAAGGGATCATCTTGGCTGCGCTGAGGAG 480
DB 480 ACCCTGCGAGCGCACAGGCTATTGGCTCAAGGGATCATCTTGGCTGCGCTGAGGAG 539
QY 481 CAGAAAGCCAAATACTTGCCTAACTGGCTCCGGGAGACATTTGCAGGCTTCTGCGCTC 540
DB 540 CAGAAAGCCAAATACTTGCCTAACTGGCTCCGGGAGACATTTGCAGGCTTCTGCGCTC 599
QY 541 ACGAGGCGAGCGAGTGGAGGATGAGCTCAATCCGAGCAGACGACACCTAAAGTAA 600
DB 600 ACGAGGCGAGCGAGTGGAGGATGAGCTCAATCCGAGCAGACGACACCTAAAGTAA 659
QY 601 GACAAGAGCACTACATCTCAATGGCTCCAGGCTGAGTAACTAATGAGGAGCTGGCC 660
DB 660 GACAAGAGCACTACATCTCAATGGCTCCAGGCTGAGTAACTAATGAGGAGCTGGCC 719
QY 661 AATATTTTACTGTTTGAAGACTGAGGTCTGATTTCTGATGATCAGTGAAGAC 720
DB 720 AATATTTTACTGTTTGAAGACTGAGGTCTGATTTCTGATGATCAGTGAAGAC 779
QY 721 AAATCAGACATTTAGTAGAAGAGACTTTTGGTGGAGTCACTAATGGAAACCCGAA 780
DB 780 AAATCAGACATTTAGTAGAAGAGACTTTTGGTGGAGTCACTAATGGAAACCCGAA 839
QY 781 GATAAATAGGCATTTCGGGCTCCAAACACTTTGCAAGTCCATTTTGAACACCAAGATA 840
DB 781 GATAAATAGGCATTTCGGGCTCCAAACACTTTGCAAGTCCATTTTGAACACCAAGATA 927

DB 840 GATAAATAGGCATTTCGGGCTCCAAACACTTTGCAAGTCCATTTTGAACACCAAGATA 899
QY 841 CCTGTGAAACATCTTTGGAGAGTTCGGAGATGGTTTAAAGTGGCCATGAACATCCTC 900
DB 900 CCTGTGAAACATCTTTGGAGAGTTCGGAGATGGTTTAAAGTGGCCATGAACATCCTC 959
QY 901 AACAGCGCCGGTTCAGCATGGGAGCGTCTGGCTGGCTGCTCAAGAGATTGATTCAA 960
DB 960 AACAGCGCCGGTTCAGCATGGGAGCGTCTGGCTGGCTGCTCAAGAGATTGATTCAA 1019
QY 961 ATGACTGCTGAGTACGCTGCACAAAGGAAACAGTTTAAACAGAGGCTCAGTGAATTTGGA 1020
DB 1020 ATGACTGCTGAGTACGCTGCACAAAGGAAACAGTTTAAACAGAGGCTCAGTGAATTTGGA 1079
QY 1021 TTGATTTCAGGAGAAATTTGCACTGATGGCTCAGAAAGGCTTACGTGATGAGAGTATGACC 1080
DB 1080 TTGATTTCAGGAGAAATTTGCACTGATGGCTCAGAAAGGCTTACGTGATGAGAGTATGACC 1139
QY 1081 TACCTCAGCAGGAGGATGCTGGACCAACCTGCTTCCGAGCTGCTCCATCGAGCAGCC 1140
DB 1140 TACCTCAGCAGGAGGATGCTGGACCAACCTGCTTCCGAGCTGCTCCATCGAGCAGCC 1199
QY 1141 ATGCTGAAGGTCTTCACTGCTCCGAGCGCTGGCAGTGTGTGAGTGAAGCGCTGCAGATC 1200
DB 1200 ATGCTGAAGGTCTTCACTGCTCCGAGCGCTGGCAGTGTGTGAGTGAAGCGCTGCAGATC 1259
QY 1201 CTCGGGGGCTTGGGCTTACAAAGGAGCTATCCGTAGGAGCGCATACTGGTGACACCCGC 1260
DB 1260 CTCGGGGGCTTGGGCTTACAAAGGAGCTATCCGTAGGAGCGCATACTGGTGACACCCGC 1319
QY 1261 ATCTCTCCATCTTCGAGGAGAACCAATGAGATTCTCCGAGTGTACATCGCTCGAGGGT 1320
DB 1320 ATCTCTCCATCTTCGAGGAGAACCAATGAGATTCTCCGAGTGTACATCGCTCGAGGGT 1379
QY 1321 CTGAGCATGCGGGCGCATCTCTGACTACAGAGATCCATGAGCTTAAACAGGCGCAAGATG 1380
DB 1380 CTGAGCATGCGGGCGCATCTCTGACTACAGAGATCCATGAGCTTAAACAGGCGCAAGATG 1439
QY 1381 AGCAGAGTATGATACCGTTCGGCGGAGCTTCGGAGCTCCCTGGCGGCAACTGTGGAC 1440
DB 1440 AGCAGAGTATGATACCGTTCGGCGGAGCTTCGGAGCTCCCTGGCGGCAACTGTGGAC 1499
QY 1441 CTGGGGCTGACAGGCAACCATGGAGTTGTGCACCCCAAGTCTTGGGAGCATGCCAACAG 1500
DB 1500 CTGGGGCTGACAGGCAACCATGGAGTTGTGCACCCCAAGTCTTGGGAGCATGCCAACAG 1559
QY 1501 TTTGAGGAGAACACCTTACTCTTTCGGCGGAGCGGTGGAGACACTGCTGCTTGGCTTGGC 1560
DB 1560 TTTGAGGAGAACACCTTACTCTTTCGGCGGAGCGGTGGAGACACTGCTGCTTGGCTTGGC 1619
QY 1561 AAGACCATCATGAGGAGGAGCTGCTTACTGAAGCGGTGGCCAAACATCCTCATCAACCTG 1620
DB 1620 AAGACCATCATGAGGAGGAGCTGCTTACTGAAGCGGTGGCCAAACATCCTCATCAACCTG 1679
QY 1621 TATGGCATGACGCGCTGCTGTCGCGGGGAGCGCTCCATCCGATTTGGGCTCCGCAAC 1680
DB 1680 TATGGCATGACGCGCTGCTGTCGCGGGGAGCGCTCCATCCGATTTGGGCTCCGCAAC 1739
QY 1681 CAGACCCAGAGGTTCTTTCGGGCAACACTTCTGCTGGGAGGCTTACTTGCAGAACTCTC 1740
DB 1740 CAGACCCAGAGGTTCTTTCGGGCAACACTTCTGCTGGGAGGCTTACTTGCAGAACTCTC 1799
QY 1741 TTCAGGCTCTCTCAGCTGGACAAAGTATGCTTCCAGAAACCTAGATGAGCAGATTAGAAA 1800
DB 1800 TTCAGGCTCTCTCAGCTGGACAAAGTATGCTTCCAGAAACCTAGATGAGCAGATTAGAAA 1859
QY 1801 GTGTCCAGCAGATTCCTTGAAGAGCGAGCTTATCTGTGCCACCCCTCTGGACAGGACA 1860
DB 1860 GTGTCCAGCAGATTCCTTGAAGAGCGAGCTTATCTGTGCCACCCCTCTGGACAGGACA 1919
QY 1861 TGC 1863
DB 1920 TGC 1922

RESULT 4

US-10-198-846-11756
: Sequence 11756, Application US/10198846
: Publication No. US2003009974A1
: GENERAL INFORMATION:
: APPLICANT: Lillie, James
: APPLICANT: Xu, Yongyao
: APPLICANT: Wang, Youzhen
: APPLICANT: Steinmann, Kathleen
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
: TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
: TITLE OF INVENTION: THERAPY OF BREAST CANCER
: FILE REFERENCE: MRI-049
: CURRENT APPLICATION NUMBER: US/10/198,846
: CURRENT FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/306,220
: PRIOR FILING DATE: 2001-07-18
: NUMBER OF SEQ ID NOS: 14084
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11756
: LENGTH: 2080
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 2075, 2076, 2077, 2078, 2079, 2080
: OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11756

Query Match 57.1%; Score 1064.6; DB 14; Length 2080;
Best Local Similarity 91.9%; Pred. No. 0;
Matches 1320; Conservative 0; Mismatches 79; Indels 38; Gaps 17;
QY 1 ATGAGCGGCTCGGGCTCTTCCTGCGCACCGGCTGCGGCTGCGCTGCGGGGTG 60
DB 491 ATGAGCGGCTCGGGCTCTTCCTGCGCACCGGCTGCGGCTGCGCTGCGGGGTG 550
QY 61 GTGGTCTCTACCGCAACCGGGCTACTGCGCACCGGCTGCGGCTGCGGCTTTCGCC 120
DB 551 GTGGTCTCTACCGCAACCGGGCTACTGCGCACCGGCTGCGGCTGCGGCTTTCGCC 610
QY 121 AAAGAGCTTTTCCTAGGCAAAATCAAGAAAGAAAGTTCCTCCATTTCCAGAGTTAGC 180
DB 611 AAAGAGCTTTTCCTAGGCAAAATCCCGAAGAG-----GTTAAGTTAGC 655
QY 181 CAAGATGAATTAATGAATCAATCAGTTCTTGGGACCGCTGGAATAATTTCTCACTGAA 240
DB 656 CCAGATGAATTAATGAATCAATCAGTTCTTGGGACCGCTGGAATAATTTCTCACTGAA 715
QY 241 GAGGTGGACTCCCGAAAAATGACACAGGAA-GGGAAAAATCCAGATGAAACTTTTGAGAA 299
DB 716 GAGGTGGACTCCCGAAAAATGACACAGGAAAGGAAATCCAGATGAAACTTTTGAGAA 775
QY 300 A-TTCAAGAGCTTAGGGCTTTTGGGCTGCG-AGTCCAGAGAAATATGGTGGCTGGGC 357
DB 776 ACTTCAAGAGCTTAGGGCTTTTGGGCTGCGAAAGTCCAGAGAAATATGGTGGCTGGGC 835
QY 358 TTCT-CCACACCATGTACTCAAGACTAGGGAGATCATCAGATGGATGGTCCATCAC 416
DB 836 TTCTCCCAACCATGTACTCAAGACTAGGGAGATCATCAGATGGATGGTCCATCAC 895
QY 417 TGTGACCTGGCAGCGCACAGGCTA-TTGGCCTCAAGGGATCATCTTGGCTGGCACTG 475
DB 896 TGTGACCTGGCAGCGCACAGGCTA-TTGGCCTCAAGGGATCATCTTGGCTGGCACTG 955
QY 476 AGGAGCAAAAGCAAAATCTTGGCTAACTTGGGCTCGGGGAGACATTCGACGCTTCT 535
DB 956 AGGAGCAAAAGCAAAATCTTGGCTAACTTGGGCTCGGGGAGACATTCGACGCTTCT 1015
QY 536 GCCTCAGGAGCCAGGCTGGGAGGATGAGGCTCAATCCGGAGCAGACCCACAC--T 593
|||||

RESULT 5

US-09-918-995-21901
: Sequence 21901, Application US/09918995
: Publication No. US20030073623A1
: GENERAL INFORMATION:
: APPLICANT: Hyseq, Inc.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
: TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
: FILE REFERENCE: 20411-756
: CURRENT APPLICATION NUMBER: US/09/918,995
: CURRENT FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: US/09/235,076
: PRIOR FILING DATE: 1999-01-20
: NUMBER OF SEQ ID NOS: 38054

DB 1016 GCCTCAGGAGCCAGCCAGTGGAGCGATGCGAG-CTCAATCCGGAGCAGAGCCACACATTT 1074
QY 594 AAGTGAAG--ACAAGAAGCAGTACATCCCTCAATGGCTCCAAGTCTCTGGATTTACTAATGGA 651
DB 1075 AAGTGAAGACACAGAGGACATACATCCCTCAATGGCTCCAGGCTCTGGGTACTAATGGA 1134
QY 652 GGACTGGCCCAATA-TTTTACTGTGTTTGCAAAGACTGAGTCTGTTGATCTGTAGTGATC 710
DB 1135 GGACTGGCCCAATA-TTTTACTGTGTTTGCAAAGACTGAGTCTGTTGATCTGTAGTGATC 1194
QY 711 AGTGAAGACAAAATCACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGCG 770
DB 1195 AGCGAAAGACCAATCCCGCGCTTCTAGTGAAGAGACTTTGGGGGGGCTCCCTAATGG 1254
QY 771 GAAACCCGGAAGATAAATTAGGCTTCGGGGCTCCAAACACTTGTGAAGTCCATTTTGAAA 830
DB 1255 GAAACCCG--GATAAATTAGGCTTCGGGGCCCCACACTTGTGTGGTCCATTTGGAGAG 1312
QY 831 CACCAAGATACCTGTGGAAACATCTTGGAGAGTCTGGAGATGGTTTAAGGTGCG-CA 889
DB 1313 CCCCCGATACCTGTGGGGAACATCTTGGGGAGGTTCGGGCCCTGGGTTCGAGGTGGCGCA 1372
QY 890 TGAACATCCTCAACAGCGCGCTTCAGCATGGCAGCGCTCGTGGCTGGGCTGCTCAAGA 949
DB 1373 TGAACATCCTCAACAGCGCGCTTCAGCATGGCAGCGCTCGTGGCTGGGCTGCTCAAGA 1432
QY 950 GATTGATTGAATGACTGCTGAGTAGCGCTGCACAAAGAAACAGTTTAAACAGAGGCTCA 1009
DB 1433 GATTGATTGAATGACTGCTGAGTAGCGCTGCACAAAGAAACAGTTTAAACAGAGGCTCA 1492
QY 1010 GTGAATTTGGATTGATTCAGGAGAAATTTGCACTGCTCAGAAAGCTTACGTCATGG 1069
DB 1493 GTGAATTTGGATTGATTCAGGAGAAATTTGCACTGCTCAGAAAGCTTACGTCATGG 1552
QY 1070 AGAGTATGACCTTACCTCAGCAGAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCA 1129
DB 1553 AGAGTATGACCTTACCTCAGCAGAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCA 1612
QY 1130 TCGAGGAGCCTGTTGAGGTGTTTCACTCGGAGGCGCGCTGGCAGTGTGTGAGTGAGG 1189
DB 1613 TCGAGGAGCCTGTTGAGGTGTTTCACTCGGAGGCGCGCTGGCAGTGTGTGAGTGAGG 1672
QY 1190 CGCTCAGATCCTCGGGGGCTTGGGCTACACAAGGAGCTATCCGTACCGAGCGCATACG 1249
DB 1673 CGCTCAGATCCTCGGGGGCTTGGGCTACACAAGGAGCTATCCGTACCGAGCGCATACG 1731
QY 1250 GTGACACCCGATCCTCCTCATCTTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCG 1309
DB 1732 GTGACACCG-ATCCTCCTCATCTTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCG 1790
QY 1310 CCCTGACGGGCTCAGCATCCGCGCGCTGACTCCTGACTACCGAGTCCATGAGCTTAAC 1369
DB 1791 CCCTGACGGGCTCAGCATCCGCGCGCTGACTCCTGACTACCGAGTCCATGAGCTTAAC 1848
QY 1370 AGGCCAAATGAGCAGCATGATGATACCGTTCGGCGGAGGCTTCGGGACTTCCCTGG 1426
DB 1849 GG---CAAGTCAGCAGCTCATCG--TACCCTTGGCGGAGGTTCGGGACTTCCCTGGG 1900

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 21901

; LENGTH: 490

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-918-995-21901

Query Match 24.3%; Score 452.2; DB 11; Length 490;

Best Local Similarity 99.3%; Pred. No. 1.9e-133;

Matches 454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1349 CCAGATCCATGAGCTTAAACAGGCCAAAGTGAGCAGCTCATGTGATACCGTTGGCCGGA 1408

DB 33 CCAGATCCATGAGCTTAAACAGGCCAAAGTGAGCAGCTCATGTGATACCGTTGGCCGGA 92

QY 1409 GGCTTCGGACTCCCTGGGCGGAACCTGTGGACCTGGGGCTGACAGCAACCATGGAGTTG 1468

DB 93 GGCTTCGGACTCCCTGGGCGGAACCTGTGGACCTGGGGCTGACAGCAACCATGGAGTTG 152

QY 1469 TGCACCCAGCTCTGGGACAGTGCCCAACAAGTTTGAGGAGAACACCTACTGCTTCGGCC 1528

DB 153 TGCACCCAGCTCTGGGACAGTGCCCAACAAGTTTGAGGAGAACACCTACTGCTTCGGCC 212

QY 1529 GGACCGTGGAGACACTGCTCTCCGCTTTGGCAAGACCATCATGAGGAGCAGCTGGTAC 1588

DB 213 GGACCGTGGAGACACTGCTCTCCGCTTTGGCAAGACCATCATGAGGAGCAGCTGGTAC 272

QY 1589 TGAAGCGGTGGCCACATCCTCATCAACCTGTATGGCATGAGCGCGTGTTCGCGG 1648

DB 273 TGAAGCGGTGGCCACATCCTCATCAACCTGTATGGCATGAGCGCGTGTTCGCGG 332

QY 1649 CCAGCGCTCCATCCGATTTGGCTCCGCAACCAACACAGGTTCTCTTTGGCCCAACA 1708

DB 333 CCAGCGCTCCATCCGATTTGGCTCCGCAACCAACACAGGTTCTCTTTGGCCCAACA 392

QY 1709 CCTTCTCGGTGGAAGCTTACTTCGAGAACTCTTTCAGCCCTCTCTCAGCTGGCAAGTATG 1768

DB 393 CCTTCTCGGTGGAAGCTTACTTCGAGAACTCTTTCAGCCCTCTCTCAGCTGGCAAGTATG 452

QY 1769 CTCAGAAAACCTTAGATGACACATTAAGAAAGTGTC 1805

DB 453 CTCAGAAAACCTTAGATGACACATTAAGAAAGTGTC 489

RESULT 6

US-09-918-995-23707

; Sequence 23707, Application US/09918995

; Publication No. US20030073623A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

; FILE OF INVENTION: FROM VARIOUS CDNA LIBRARIES

; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995

; PRIOR FILING DATE: 2001-07-30

; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 23707

; LENGTH: 483

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-918-995-23707

Query Match 24.3%; Score 451.8; DB 11; Length 483;

Best Local Similarity 99.6%; Pred. No. 2.5e-133;

Matches 453; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1349 CCAGATCCATGAGCTTAAACAGGCCAAAGTGAGCAGCTCATGTGATACCGTTGGCCGGA 1408

DB 29 CCAGATCCATGAGCTTAAACAGGCCAAAGTGAGCAGCTCATGTGATACCGTTGGCCGGA 88

QY 1409 GCCTTCGGACTCCCTGGGCGGAACCTGTGGACCTGGGCTGTACAGCAACCATGGAGTTG 1468

DB 89 GCCTTCGGACTCCCTGGGCGGAACCTGTGGACCTGGGCTGTACAGCAACCATGGAGTTG 148

QY 1469 TGCACCCAGCTTTCGGGACAGTGCCCAACAAGTTTGAGGAGAACACCTACTCTCTTCGGCC 1528

DB 149 TGCACCCAGCTTTCGGGACAGTGCCCAACAAGTTTGAGGAGAACACCTACTCTCTTCGGCC 208

QY 1529 GGACCGTGGAGACACTGCTCTCCGCTTTGGCAAGACCATCATGAGGAGCAGCTGGTAC 1588

DB 209 GGACCGTGGAGACACTGCTCTCCGCTTTGGCAAGACCATCATGAGGAGCAGCTGGTAC 268

QY 1589 TGAAGCGGTGGCCCAACATCCTCATCAACCTGTATGGCATGAGCGCGTGTTCGCGG 1648

DB 269 TGAAGCGGTGGCCCAACATCCTCATCAACCTGTATGGCATGAGCGCGTGTTCGCGG 328

QY 1649 CCAGCGCTCCATCCGATTTGGCTCCGCAACCAACACAGGTTCTCTTTGGCCCAACA 1708

DB 329 CCAGCGCTCCATCCGATTTGGCTCCGCAACCAACACAGGTTCTCTTTGGCCCAACA 388

QY 1709 CCTTCTCGGTGGAAGCTTACTTCGAGAACTCTTTCAGCCCTCTCTCAGCTGGCAAGTATG 1768

DB 389 CCTTCTCGGTGGAAGCTTACTTCGAGAACTCTTTCAGCCCTCTCTCAGCTGGCAAGTATG 448

QY 1769 CTCAGAAAACCTTAGATGACAGCATTAAGAAAGTG 1803

DB 449 CTCAGAAAACCTTAGATGACAGCATTAAGAAAGTG 483

RESULT 7

US-09-918-995-11034

; Sequence 11034, Application US/09918995

; Publication No. US20030073623A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

; FILE OF INVENTION: FROM VARIOUS CDNA LIBRARIES

; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995

; PRIOR FILING DATE: 2001-07-30

; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 11034

; LENGTH: 451

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-918-995-11034

Query Match 21.0%; Score 390.6; DB 11; Length 451;

Best Local Similarity 97.8%; Pred. No. 8.3e-114;

Matches 396; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 585 AGCCACACTAAGTGAAGACAAAGCAGCTACATCTCAATGGCTCCCAAGCTCTGGATTAC 644

DB 47 AGCCACACTAAGTGAAGACATGAAGCAGCTACATCTCAATGGCTCCCAAGCTCTGGATTAC 106

QY 645 TAATGGAGGACTGGCCAAATATTTTACTGTGTGTGCAAGACACTGAGGTCGTTGATTCTGA 704

DB 107 TAATGGAGGACTGGCCAAATATTTTACTGTGTGTGCAAGACACTGAGGTCGTTGATTCTGA 166

QY 705 TGGATCAGTGAAGACAAAATCACAGCATTCATAGTAGAAGAGACTTTTGGTGGAGTCAC 764

DB 167 TGGATCAGTGAAGACAAAATCACAGCATTCATAGTAGAAGAGACTTTTGGTGGAGTCAC 226

QY 765 TAATGGAAACCCCAAGATAAATTAGGCATTCGGGGCTCCCAACACTTGTGAAGTCCATTT 824

DB 227 TAATGGAAACCCCAAGATAAATTAGGCATTCGGGGCTCCCAACACTTGTGAAGTCCATTT 286

QY 825 TGAACACCAAGATACCTCTGTGAAACATCTTGGAGAGCTCGGAGATGGGTTTAAGGT 884

DB 287 TGAATACCAAGATACCTCTGTGAAACATCTTGGAGAGCTCGGAGATGGGTTTAAGGT 346

Qy	885	GGCCATTGAACATCTCTCAACAGCGGCCGTTTCAGCATGGCAGCTCGTGGCTGGGCTGCT	944
Db	347	GGTCATTGAACATCTCTCAACAGCGGCCGTTTCAGCATGGCAGCTCTTGGCTGGGCTGCT	406
Qy	945	CAAGAGATTGATTGAAATGACTGCTGAGTACGCCCTGCACAAGGAA	989
Db	407	CAAGAGATTGATTGAAATGACTGCTGAGTACGCCCTGCACAAGGAA	451

RESULT 8

```

US-09-917-800A-1584
; Sequence 1584, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1584
; LENGTH: 2117
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012891
US-09-917-800A-1584

```

Query Match	20.9%	Score 388.6;	DB 10;	Length 2117;
Best Local Similarity	54.5%;	Misid. No. 9.e-113;		
Matches 891;	Conservative 0;	Primal. Noches 729;	Indels 15;	Gaps 5;
113 QY	CTTTCCGCAAGAGCTTTTCC	TAGGCANAATCAAGAAAGAAGT	TTTCCCATTTCCAG	172
Db				
236 QY	CTTTTGTGTGGNATGTTTCA	AGGCCAGCTTACCACCGCAGGT	GTGTCCCATACCAT	295
Db				
173 QY	AGTTTAGCCNAGATGAAC	TTTAATGAA---ATCAATCAG	TTCTTGGGACCCGTGGAAAAT	229
Db				
296 QY	CTGTGCTCAATGAAGGAC	AGACACAATTTCTCAAGAG	CGTGTGGGACGTCGGCCGGT	355
Db				
230 QY	TCTTCACTGAAGAGGTGG	ACTCCCGAAAAAATTGACC	AGGAGGAATCCAGATGAA	289
Db				
356 QY	TCCTTTGAAGAAGTGA	TGACCCCTGCCAAGATG	ACTCCTTGGAGAGGTGGAGAGGACA	415
Db				
290 QY	CTTTGGAGAAATTTGAG	AGCCTTAGGCTTTTTTGG	CTGCAAGTCCCGAAGAATATGGT	349
Db				
416 QY	CTTTGCAAGGACATCA	AAGAACTATGGGGCAATTTGGT	CTGCAAGTACCCAGCAGCTGGGT	475
Db				

QY	350	GCCTGGCCTTCTCCAAACACCAATGTACTCAAGACTAGGGAGATCATACGACATGGATG---	400
Db	476	GTTTGGGCGCTCTCTAATACCCATGACGCTCGCTTGCACAGATCTGTGGCAGTGCATGACC	535
QY	407	GGTCCATCACTGTGACCCCTGGCAGGCGCACAGGCTATTGGCCCTCAAGGGGATCATCTTGG	466
Db	536	TTGGTGTTAGCGTTACCTCTGGGAGCCATCAGAGCATCGGTTTCAAGGGCATCTTGGCTCT	595
QY	467	CTGGGCACTCAGGAGCAGAAAGCAAAATACTTGCCTTAAACTTGGCGTCCGGGAGCACATTTG	526
Db	596	ATGGCACAAGGCCCAAGAGGAAAAATACCTCCCAAGATGSCATCCGGCAGGCTTTGG	655
QY	527	CAGCCTTTCGCTCAGGGACGACGCTGGGAGCGATGCCAGCCTCAATCCGGAGCACAG	586
Db	656	CGGCTTTCGCTGACTGAGCCCTCGAGGGGTCCGATGTGGCCCTTATTTTCTTCTTCTT	715
QY	587	CCACACTAAGTGAAGACAAAGACACTACATCTCTCAATGGCTTCCAAAGTCTGGATTACTA	646
Db	716	CTGTACTTAGCCCTGTGAAAAGTATTATCTCTCAACGGAAGCAAGATCTTGGATCAGTA	775
QY	647	ATGGAGGACTGCCAATATTTTACTGTGTTTGGAAAGACTGAGGTCGTTGATCTGTCTG---	703
Db	776	ATGGGGGTCTGCACACATTTTCACTGTCTTTTGGCAAAACGCCAATTAAGATCGAGCCA	835
QY	704	ATGGATCATGTGAAGACAAAAATCACAGCATTCATAGTAGAAAGAGACTTTGGTGGAGTCA	763
Db	836	CGGGGCGGTGAAGACAGAGATCACAGCTTTCGTAGTGAACGGAGCTTTGCAGGGGTTA	895
QY	764	CTAATTTGGAAACCCGAAGATTAATTTAGGCATTCGGGGCTCCAACACTTTGTGAAGTCCATT	823
Db	896	CCCATTTGGGCTCCCCGAAAGAGATGGGCATCAAGGCATCTTAACACATCAGAGGTGACT	955
QY	824	TTGAAAACACCAAGATACCTGTGGAACACATCCTTTGGAGAGGTCGGAGATGGGTTTAAAG	883
Db	956	TTGATTTGGAGTCAAGGTGCCAGCAGAAATGTCTAGGAAAGTGGGAGATGGCTTTCAAGG	1015
QY	884	TGGCCATGAACATCCTCAACAGCGCCGGTTTTCAGCATGGGCAGCGTCTGGCTGGGGTGC	943
Db	1016	TTGCTGTCAACATCTCAACACGGAAGATTTGGGATGGCTTGAACCTAGCAGGCACCA	1075
QY	944	TCAAGAGATTGATTGAATGACTCTGTAGTAGCCTGCACAAAGCAACAGTTTAAACAAGA	1003
Db	1076	TGAAGCCATCATTTGCCAAGCGGTTTATCATGTCTACTAACCGTACCACGATTTGGGGACA	1135
QY	1004	GGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCACGTATGGCTCAGAAGGCTTACG	1063
Db	1136	AAATTCACAACCTTTGGGTGATTCAGGAAAGCTGGCTCGGATGGCTATTCTGCAGTATG	1195
QY	1064	TCATGGAGAGTATGACCTACCTCACAGCGGGATCTGGACCAACCTTGGCTTCCCGACT	1123
Db	1196	TGACTGAGTCCATGGCTTTACATGCTGAGTGCCAAACATGGACCAAGTGGAGCTTCAAG	1252
QY	1124	GCTCCATCGAGGAGCCATGTGTGAAGGTGTTTTCAGCTCCGAGGCGGCTGGCAGTGTGTGA	1183
Db	1253	TCCAGATAGAAGCTGCCATCAGCAAAATCTTTGGCTCGGAGCGGCTCGAAAGTGCAG	1312
QY	1184	GTGAGGCGCTGCACATCTCTGGGGGCTTGGCTTACACAAGGGACTATCCCTCAGACGGCA	1243
Db	1313	ATGATGCATCCAGATTAATGGGGGCTGGGCTTTCATGAAGAACCAAGGGGTAGAGCGTG	1372
QY	1244	TACTTGGCTGACACCCGCTCTCTCTCATCTTTTCAGGGAACCAATGAGATTCTCCGGATGT	1303
Db	1373	TGCTCCGAGATATTCGAATCTTCCGGATCTTTGAGGGACAAATGACATTTCTTTCGACTGT	1432
QY	1304	ACATCGGCCCTGACGGGTCTCGAGCATGCCGGCCGCATCTCTGACTACACAGGATCCATGAGC	1363
Db	1433	TTTGTGGCTCTACAAGGCTGCATGGACAAAGAAAGGAACCTCACGGGACTTGGTAAATGCC	1492
QY	1364	TTAAACAGGCCAAAGTCAGCACAGTCATGATACCGTTGGCCGGAGGCTTCGGGACTCCC	1423
Db	1493	TAAAGATTCCTTTTGGAAATGTTGGCCCTCTCTATAGGAGAAAGCAACAAACAGCTGAGC	1552
QY	1424	TGGGCCGAAGCTGTGAGCTTGGGCTCGACGGCAACCATGAGTGTGTGGACCCCGAGTCTTG	1483

Db 1553 GCGGACAGGATTGGCACTGGTCTGA---GTCTCTCGGAATTGTCCACCCAGAGTTGA 1609
QY 1484 CGGACAGTGCCAAACAGTTTGAGGAGACACCTACTGCTTGGCCGGACCGTGGAGACAC 1543
Db 1610 GTCCGAGTGGTGAACCTGGCAGTGCAGGCTCTGGAACAATTTGCCACTGTAGTGGAGGCGA 1669
QY 1544 TGCTGCTCGCTTTGGCAGACCATCATGGAGGAGCAGTGTACTGAAGCGGGTGGCCA 1603
Db 1670 AGCTGATGAAGCACAAAGAGGATGTCAATGAACAGTTCTGTGTGAGGACTGGCAG 1729
QY 1604 ACATCCTCATCAACCTGTATGCATGACGGCGTGTGTGCGGGCCAGCCGCTCCATCC 1663
Db 1730 ATGGAGCCATTGACCTTAGCCATGGTGTGTTCTCTCCAGAGCCTCAAGATCCCTGA 1789
QY 1664 GAATGGGCTCGGCAACAGCAGCAGGAGTCTCTTGGCCAAACCTTCTCGTGGAG 1723
Db 1790 GTGAGGGCTACCCGACAGCAGCATGAGAAAATGCTGTGATAGTTGGTGCATTGAGG 1849
QY 1724 CTTACTTGCAGATC 1738
Db 1850 CTGCAACAGGATTTC 1864

RESULT 9

US-09-918-995-16927
; Sequence 16927, Application US/0918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16927
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-16927

Query Match 14.08; Score 261.4; DB 11; Length 490;
Best Local Similarity 99.3%; Pred. No. 1.6e-72;
Matches 273; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1560 CAAGCATTATGAGGAGCAGCTGGTACTGAAGCGGGTGGCCAAACATCTCATCAACCT 1619
Db 217 CCAGACCATATGAGGAGCAGCTGGTACTGAAGCGGGTGGCCAAACATCTCATCAACCT 276
QY 1620 GTATGGCATGACGGCGGTGTGTGCGGGCCAGCCGCTCCATCCGATTGGGTCCGCAA 1679
Db 277 GTATGGCATGACGGCGGTGTGTGCGGGCCAGCCGCTCCATCCGATTGGGTCCGCAA 336
QY 1680 CCAGGACAGAGTTCTTCTTGGCCAAACCTTCTCGGTGGAAGCTTACTTGCAGATCT 1739
Db 337 CCAGGACAGAGTTCTTCTTGGCCAAACCTTCTCGGTGGAAGCTTACTTGCAGATCT 396
QY 1740 CTTGAGCTCTCTCAGTGGACAGTATGCTCCAGAAAACCTAGATGAGCAGATTAAAGAA 1799
Db 397 CTTGAGCTCTCTCAGTGGACAGTATGCTCCAGAAAACCTA-ATGAGCAGATTAAAGAA 455
QY 1800 AGTGTCCAGCAGATCCTTGAGAGCGAGCCTATA 1834
Db 456 AGTGTCCAGCAGATCCTTGAGAGCGAGCCTATA 490

RESULT 10

US-09-822-849A-480
; Sequence 480, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakr
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822,849A
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 480
; LENGTH: 312
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-480

Query Match 11.5%; Score 214.6; DB 9; Length 312;
Best Local Similarity 96.1%; Pred. No. 1e-57;
Matches 220; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ATGAGCGGTGCGGGCTCTTCTGCGCACCGGTGGGCTGCGTGGTGGCGGGGTCTG 60
Db 57 ATGAGCGGTGCGGGCTCTTCTGCGCACCGGTGGGCTGCGTGGTGGCGGGGTCTG 116
QY 61 GTGGTCTCTACCGCGAAGCGGGTACTGCGCACCGGTGGGCTGCTAGAGCTTTGCGC 120
Db 117 GTGGTCTCTACCGCGAAGCGGGTACTGCGCACCGGTGGGCTGCTAGAGCTTTGCGC 176
QY 121 AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAGTTTCCCATTTCCAGAAAGTTAGC 180
Db 177 AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAGTTTCCCATTTCCAGAAAGTTAGC 236
QY 181 CAAGATGAATTAATGAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 229
Db 237 CAAGATGAATTAATGAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 285

RESULT 11

US-09-796-692-8175
; Sequence 8175, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28

```

; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8175
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8175

```

```

Query Match      8.9%; Score 165.2; DB 10; Length 568;
Best Local Similarity 58.0%; Pred. No. 9.2e-42;
Matches 312; Conservative 0; Mismatches 223; Indels 3; Gaps 1;

QY 782 ATAAATTAGGCATTCGGGGCTCCCAACACTTCTGGAAGTCCATTTTGAAACACCAAGATAC 841
Db 3 AGAAGATGGGCATCAAGGCTTCAACACACAGCAGAGGTGTTCTTGTATGGAGTACGGGTGC 62

QY 842 CTGTGGAAAACATCCTTGGAGAGGTCGGAGATGGGTTTAAAGTGCCCATTAACATCCTCA 901
Db 63 CATCGGAGAACGTCGTGGGTGAGGTGGGAGTGGCTTCAAGGTTGCCATGCACATCCTCA 122

QY 902 ACAGCGCGGCTTCAGCATGGCAGCGCTGCTGGCTGGCTGCTCAAGAGATTGATGAA 961
Db 123 ACATGGAGGTTTGGCATGGCTGGGCGCTTGGCAGGTACCATGAGAGGCATCTTGCTA 182

QY 962 TGACTGCTGAGTACGCTTCACAAGAAACAGTTTAAACAGAGGCTCAGTGAATTTGGAT 1021
Db 183 AGCGGTAGATGCTTAACTATGAGCTTGGGAGAGAAATTCACAACATTTGGGC 242

QY 1022 TGGTGAAGGTTTTCAGTCCGAGGCGCGCTTGGCAGTGTGTGAGTCAGGCGCTGCAGATCC 1201
Db 360 TCAGCAAAATCTTTGGCTCGGAGGCGAGCTGGAAGGTGACAGATGAATGCATCCAAATCA 419

QY 1202 TCAGGAGGTTGAGGCTGACAGAGGCAATGAGATTTCTCCGGATGTACATCCCTGACGGG 1319
Db 480 TCCTCCGATCTTTGAGGGGCAATGACATCTTCGGCTGTTTGGCTCTGCGAGG 537

```

RESULT 12

```

US-10-040-862-8175
; Sequence 8175, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Aigate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation

```

```

; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8175
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-8175

```

```

Query Match      8.9%; Score 165.2; DB 14; Length 568;
Best Local Similarity 58.0%; Pred. No. 9.2e-42;
Matches 312; Conservative 0; Mismatches 223; Indels 3; Gaps 1;

QY 782 ATAAATTAGGCATTCGGGGCTCCCAACACTTCTGGAAGTCCATTTTGAAACACCAAGATAC 841
Db 3 AGAAGATGGGCATCAAGGCTTCAACACACAGCAGAGGTGTTCTTGTATGGAGTACGGGTGC 62

QY 842 CTGTGGAAAACATCCTTGGAGAGGTCGGAGATGGGTTTAAAGTGCCCATTAACATCCTCA 901
Db 63 CATCGGAGAACGTCGTGGGTGAGGTGGGAGTGGCTTCAAGGTTGCCATGCACATCCTCA 122

QY 902 ACAGCGCGGCTTCAGCATGGCAGCGCTGCTGGCTGGCTGCTCAAGAGATTGATGAA 961
Db 123 ACATGGAGGTTTGGCATGGCTGGGCGCTTGGCAGGTACCATGAGAGGCATCTTGCTA 182

QY 962 TGACTGCTGAGTACGCTTCACAAGAAACAGTTTAAACAGAGGCTCAGTGAATTTGGAT 1021
Db 183 AGCGGTAGATGCTTAACTATGAGCTTGGGAGAGAAATTCACAACATTTGGGC 242

QY 1022 TGGTGAAGGTTTTCAGTCCGAGGCGCGCTTGGCAGTGTGTGAGTCAGGCGCTGCAGATCC 1081
Db 243 TCATCCAGGAGAGCTGCGCAGGATGGTTATGCTGCAAGTATGTAACATGATGCTTGGCTT 302

QY 1082 ACCTCAGCAGCAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCCGAGCGCA 1141
Db 303 ACATGGTGAAGTCTTAACTATGAGCTTGGGAGAGAAATTCACAACATTTGGGC 359

QY 1142 TGGTGAAGGTTTTCAGTCCGAGGCGCGCTTGGCAGTGTGTGAGTCAGGCGCTGCAGATCC 1201
Db 360 TCAGCAAAATCTTTGGCTCGGAGGCGAGCTGGAAGGTGACAGATGAATGCATCCAAATCA 419

QY 1202 TCAGGAGGTTGAGGCTTACAGAGGCAATGAGATTTCTCCGGATGTACATCCCTGACGGG 1261
Db 480 TCCTCCGATCTTTGAGGGGCAATGACATCTTCGGCTGTTTGGCTCTGCGAGG 537

```


Qy	1235	ACGAGCGCATAC	TGCGTGTAC	CCCGATCCT	CTCATCTT	TCGAGGAA	CCAATGAGATTC	1294
Db	1061	TGCCCCGGATGT	GGCGCGATCC	AAGATCCT	TGGAGATCG	CGAGGAC	CACGCGAGGTGC	1120
Qy	1295	TCCGGATGTAC	ATCGC	1310				
Db	1121	AGCGGATGCT	GATCGC	1136				

RESULT 15

```

US/10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IREDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156.761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or
US/10-156-761-1

```

Query Match	8.1%;	Score 151.2;	DB 14;	Length 9025608;
Best Local Similarity	49.3%;	Pred. No. 1.4e-34;		
Matches 511;	Conservative 0;	Mismatches 513;	Indels 12;	Gaps 4;
QY	278	TCCAGATCAAACCTTTGGAGAAATTCAGAGCCTAGGGCTTTTGGCTGCCAGTCCCAG	337	
Db	6391537	TCCCGTACGAGATCGTCCCGGAGATGGCGGATGGCCGTGTCGGGCTGCCCTGCCG	6391478	
QY	338	AGAAATATGGTGGCTTGGCTTCTCCAA---CACCATGTACTCAAGACTAGGGGAGATCA	394	
Db	6391477	AGAGTACGGCGGCATGGCGCGGACTATCTGCGCTCGGCATCGCCCTCGAAGAAGCTCG	6391418	
QY	395	TCAGCATGATGGGTCCATCACTGTGACCCCTGGCAGCGCACAGGCTATTGGCCTCAAGG	454	
Db	6391417	CCCGGTGCACTCTCCGTGCGCATCACCCCTGGAGCCGGGGTCTCAGTGGGCGCATGC	6391358	
QY	455	GAATCATCTTGGCTGGCACTGAGGAGCAGAAAGCCAAATACTTGCCTAAACTGGCGTCCG	514	
Db	6391357	CGATCCAGCTCTTCGGGACGGAGCGCAGAGGCGGAGTGGCTCCCCCGGCTGTGTTCCG	6391298	
QY	515	GGGAGCACATGCAAGCCTTCTCCCTCAGGAGCCAGCCAGTGGGACCGATGCAGCCTCAA	574	
Db	6391297	CGGAGATCTCGGCGCCTTCGGTCTCACCAGCGGACGGCGGCTCGGACGGCGGGCGCA	6391238	
QY	575	TCCGGAGCAGAGCCACACTAAGTGAAGACAAAGACACTACATCTCTCAATGGCTCCAAAGG	634	
Db	6391237	CGCGCACAGCGCCCGCCTGGACGAGTGCAGCAACGAAATGGGTGATCAACGCGCAACAAT	6391178	
QY	635	TCTGGATTACTAATGAGGACTGGCCATATTTTACTGTGTTTGCAAAGACTGAGGTGCG	694	
Db	6391177	GCTTCATCCCAACTGGGCA---CCGACATCAGCGGGTGTGTGACGTCACGGGGGTCA	6391121	

Qy	695	TTGATTTCTGATGGATCAGTGAAGACAAAAATCACAGCATTCATAGTAGAAGAGACTTTG	754
Db	6391120	CCGCCCGGAAGCCGACGGGAACCGGTGATCTCTCGATCATGCTCCGTCGGCACGC	6391061
Qy	755	GTGCAGTCACTAATGTGGAAACCCGAAGATAAATTAGGCAATTCGGGGCTCCAACACTTGTG	814
Db	6391060	CCGGCTTTCAGGTGCGCGCCCGGTACTCGAAGGTTCGGCTGGAACGCCTCGGACACCCGCG	6391001
Qy	815	AAGTCCATTTTGAACACCAAGATACCTGTGAAACACATCTTTGGAGAGGTTCGGAGATG	874
Db	6391000	AGCTGTCTTTCGCGGAGCTCCGGTCCCGCCGCGAACCTGCTGGCGGAACAGGCCGCG	6390941
Qy	875	GGTTTAAAGTTCGCCATCAACATCCTCAACAGCGCGGTTTCAGCATGGGCAGGTCGTGG	934
Db	6390940	GGTACGCGCATTTCTTCGGATCCTTCACAGGAGCGGATTCGCCATCTCGGCCCTGGCCA	6390881
Qy	935	CTGGCTGCTCAAGAGATTGATTTGAATGACTGCTGAGTAGCGCTGCACAAAGAAACAGT	994
Db	6390880	CCGGCTCGCCAGGGCTGCGTGGACGAGTCGCTGAAGTACGCCCGGGGAACGCCACGCGT	6390821
Qy	995	TTACAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCACCTGATGGCTCAGA	1054
Db	6390820	TCGGCGGGAACATCGGGCGGTACACAGGCCATCCAGTTCAGATCGCCGATATGGAGATGA	6390761
Qy	1055	AGCTTACGTCATGGAGAGTATACCTACCTCACAGCAGGATGCTGGACCAACCTGGCT	1114
Db	6390760	AGCGCGA---CATGGCCCGCTCGGTGGCGTGACCGCCGCTCCGCCCTGGTCGCCGGCG	6390704
Qy	1115	TTCCCACTGCTCCATCGAGGCACCCATGGTGAAGGTGTTCAGCTCCGAGGCGCCCTGGC	1174
Db	6390703	AGCC---CTTCAAGAAGGAGCGCGCATCGCCAAAGCTCTACTCGTCCACGGTCCGCTCG	6390647
Qy	1175	AGTGTGTGAGTGAGGCGCTGCAGATCCTCGGGGGCTTGGGCTACACAAGGACATATCCGT	1234
Db	6390646	ACAACCCCGGAGGCCACCCAGATCCAGCGCGGCTTACGGCTTCATGAACAGATACCCCG	6390587
Qy	1235	ACGAGCGCTACTTCGCTGACACCCGATCCTCTCATCTTCAGAGGAACCAATGAGATTC	1294
Db	6390586	TGCCCCGGATGTGCCGCACTCCAAGATCCTGGAGATCGCGGAGGACCAACGAGGTGC	6390527
Qy	1295	TCCGGATGTACATCGC	1310
Db	6390526	AGCGGATGCTGATCGC	6390511

Search completed: September 7, 2003, 05:59:04
Job time : 432.388 secs

THIS PAGE BLANK (USPTO)